

Untersuchungen zu psychoakustischen Mess- und Berechnungsverfahren der Tonhaltigkeit

Von der Fakultät für Mathematik und Naturwissenschaften
der Carl von Ossietzky Universität Oldenburg
zur Erlangung des Grades und Titels
eines Doktors der Naturwissenschaften (Dr. rer. nat.)
angenommene Dissertation

von Herrn Matthias Vormann
geboren am 20.07.1968
in Hagen/Westfalen

Gutachter: Prof. Dr. Jesko Verhey
Zweitgutachter: Prof. Dr. Volker Mellert
Tag der Disputation: 20.05.2011

Zusammenfassung

Umweltschalle werden generell als störender empfunden, wenn sie deutlich aus dem Hintergrundschaall herausragende, tonale Anteile beinhalten. Daher kommt dieser als Tonhaltigkeit bezeichneten Geräuscheigenschaft eine wichtige Rolle im Lärmschutz bei der Beurteilung von Geräuschmissionen zu.

Die Tonhaltigkeit wird in der Norm DIN 45681 („Akustik - Bestimmung der Tonhaltigkeit von Geräuschen und Ermittlung eines Tonzuschlages für die Beurteilung von Geräuschmissionen“) unter messtechnischen Gesichtspunkten definiert. Es werden die zu verwendenden objektiven Verfahren zur Ermittlung der Tonhaltigkeit festgelegt und ihr Anwendungsbereich dargelegt. Die Wahrnehmung der Tonhaltigkeit wird jedoch nicht ausreichend thematisiert.

Das Ziel der vorliegenden Arbeit ist die Untersuchung der Wahrnehmung von Tonhaltigkeit und der zugrunde liegenden Mechanismen und Phänomene. Dazu wird im ersten Schritt das Repertory-Grid Verfahren nach Kelly als ein geeignetes Instrument zur Untersuchung der wahrgenommenen Tonhaltigkeit identifiziert. Mit diesem Verfahren wird die interne Repräsentation der Tonhaltigkeit untersucht und daraus der Wahrnehmungsraum abgeleitet und diskutiert.

Im zweiten Schritt wird vor diesem Hintergrund ein adäquater Orientierungsschaall für die Demonstration der Tonhaltigkeit erarbeitet.

Der dritte Schritt umfasst die Entwicklung eines universellen Messverfahrens zur subjektiven Erfassung der Tonhaltigkeit beliebiger Schalle. Die auf diese Weise gewonnenen Ergebnisse sind direkt interpretierbar und erlauben so unmittelbar quantitative Aussagen zur Tonhaltigkeit.

Im vierten Schritt werden die für die Tonhaltigkeit relevanten Parameter untersucht. Die bisher in der Norm nicht angemessen berücksichtigten Einflußgrößen werden zur Ergänzung der DIN 45681 spezifiziert und modelliert.

Abstract

Environmental sounds often contain tonal components which are commonly embedded in a noise floor. These tonal components are usually considered as annoying and therefore play an important role in the assessment of environmental noise.

The measurement and evaluation of the magnitude of tonal content is defined in the German standard DIN 45681 („Acoustics - Determination of tonal components of noise and determination of a tone adjustment for the assessment of noise immissions”). However, the standard is mainly concerned with the assessment of the annoyance and does not clearly define the sensation that is elicited by tonal content.

The aim of this thesis is to reveal the perception of tonal components in a noise floor. In a first step, the Repertory Grid method, established by Kelly, is identified as an adequate instrument for the investigation of the perception of this tonal content. With this method, the magnitude of tonal content is analyzed and the internal representation (perceptual space) is evaluated and described.

In a second step, a suitable orientation stimulus is developed to familiarize listeners with the sensation of tonal components.

In a third step, a method for measuring the perceived magnitude of tonal content of any kind of signals and noises is developed and evaluated. The results of these experiments can be directly interpreted in the light of the standard.

It is shown that several stimulus parameters affect the magnitude of tonal content. An extension of the standard is suggested to include these effects for the estimation of the magnitude of tonal content.

Inhaltsverzeichnis

1. Einleitung und Motivation	1
2. DIN 45681	3
2.1. Messdurchführung	3
2.2. Eingangswerte der THK-Berechnung	3
2.3. Maßgebliche Differenz ΔL	4
2.4. Tonzuschlag K_T	4
2.5. Frequenzabhängigkeit der THK	5
2.6. Ausprägtheit von Bandpassrauschen	5
2.7. Genau zwei Töne in einer Frequenzgruppe	6
2.8. Mehr als zwei Töne in einer Frequenzgruppe	6
2.9. Töne in unterschiedlichen Frequenzgruppen	6
2.10. Nichtstationäre Signale	7
2.11. Probleme bei der Bestimmung des Schmalbandpegels	7
3. Perzeptuelle Aspekte der Tonhaltigkeit	9
3.1. Repertory-Grid Verfahren	12
3.1.1. Grundlagen	12
3.1.2. Methode	13
3.2. Setup der Messungen	18
3.2.1. Probanden	18
3.2.2. Durchführung	18
3.3. Untersuchung der maßgeblichen Differenz	19
3.3.1. Stimuli	19
3.3.2. Evokationsphase: Auswertung der Konstrukte	21
3.3.3. Evokationsphase: Auswertung durch MDS	23
3.3.4. Rangordnung der Elemente (Stimuli)	27
3.3.5. Skalierung der Elemente (Stimuli)	29
3.4. Diskussion	31
4. Wahl eines Orientierungsschalls für die THK	33
4.1. Motivation und Ziel der Untersuchung	33
4.2. Probanden	33
4.3. Stimuli	34

4.4. Durchführung	36
4.5. Bewertungsergebnisse	37
4.6. Diskussion	43
4.7. Standard-Orientierung für die Tonhaltigkeit	44
5. Tonhaltigkeit von Einzeltönen in Rauschen im vollständigen Paarvergleich	49
5.1. Motivation und Ziel der Untersuchung	49
5.2. Probanden	49
5.3. Durchführung	50
5.4. Stimuli	50
5.5. Abhängigkeit der THK von der maßgeblichen Differenz	52
5.6. Abhängigkeit der THK von der Frequenz	57
5.7. Diskussion	60
6. Einfluss verschiedener Signalparameter auf die Tonhaltigkeit	63
6.1. Motivation und Ziel der Untersuchung	63
6.2. Adaptives THK-Messverfahren	64
6.2.1. Verfahren	64
6.2.2. THK Referenzsignal	66
6.3. Methode	72
6.3.1. Probanden	72
6.3.2. Anwendung und Konfiguration des Verfahrens	72
6.3.3. Darstellung der Ergebnisse	74
6.4. Messungen	76
6.4.1. THK als Funktion der Anzahl von Tönen	76
6.4.2. THK als Funktion der Harmonizität von Teiltönen	81
6.4.3. THK als Funktion des Abstandes zweier Sinustöne	84
6.4.4. THK als Funktion des absoluten Darbietungspegels	86
6.4.5. THK als Funktion der Bandbreite von Schmalbandrauschen	88
6.4.6. THK als Funktion der Modulationsfrequenz (AM) modulierter Töne	91
6.4.7. THK als Funktion der Modulationsfrequenz (FM) modulierter Töne	93
6.4.8. THK als Funktion der Tonfrequenz	94
6.4.9. THK als Funktion der Tonfrequenz II	96
6.5. Zusammenfassung	100
7. Verfahrensvergleich zur Bestimmung der Tonhaltigkeit	103
7.1. Stimuli und Messverfahren	104
7.2. Signaldarbietung	105
7.3. Probanden	106

7.4. Mithörschwelle für die HKTs	106
7.4.1. Verfahren	106
7.4.2. Ergebnisse	107
7.4.3. Diskussion	108
7.5. Variation der Verfahren zur Messung der THK	110
7.6. Ergebnisse des Verfahrensvergleich	112
7.6.1. Bewertung der HKTs als Teststimuli	113
7.6.2. Bewertung der HKTs als Vergleichsstimuli	115
7.6.3. Vergleich der Bewertung als Vergleichs- bzw. Teststimuli	118
7.6.4. Vergleich der Messdesigns ET und IT sowie der Startpegel	119
7.6.5. Zusammenfassung und Diskussion der Messungen 1–8	121
7.6.6. Bewertung der HKTs mit Blocked Tracks	123
7.7. Modellierung der THK für HKTs	124
7.8. Diskussion	128
8. Zusammenfassung	133
Anhänge	135
A. Breitbandiges Rauschen, Verdeckung und Verdeckungsmaß	135
A.1. Breitbandiges Rauschen	135
A.2. Verdeckungsmaß	137
B. Geräusch-Sets und Signale	139
B.1. Pompetzki	139
C. Perzeptuelle Aspekte der Tonhaltigkeit	141
C.1. Evokationsphase	141
C.2. Beurteilungsphase	141
D. Wahl eines Orientierungsschalls für die THK	145
D.1. Orientierungsschalle	145
D.2. Tonhaltigkeitsverlauf der Orientierungsschalle	145
D.3. Individuelle Probandenurteile	147
D.4. Law of Categorical Judgement	147
D.5. Erweiterte Analyse der Daten (ANOVA)	150
E. THK von Einzeltönen in Rauschen im vollständigen Paarvergleich	151
E.1. Instruktion der Probanden im Wortlaut	151
E.2. Berechnete Tonhaltigkeit der verwendeten Stimuli	151
F. Einfluss verschiedener Parameter auf die THK	155
F.1. Verwendete Stimuli	155

F.2. Verdeckungsmaß für die Teiltöne der HKTs	156
F.3. Erweiterte Datenanalyse (ANOVA)	156
F.3.1. Anzahl von Tönen	158
F.3.2. Harmonizität von Teiltönen	159
F.3.3. Abstand zweier Sinustöne	159
F.3.4. Absoluter Darbietungspegel	160
F.3.5. Bandbreite von Schmalbandrauschen	162
F.3.6. Modulationsfrequenz (AM) modulierter Töne	163
F.3.7. Modulationsfrequenz (FM) modulierter Töne	164
F.3.8. Tonfrequenz	165
G. Verfahrensvergleich	169
G.1. Normierung der Vergleichsstimuli	169
G.2. Transformation der Messwerte für Blocked Tracks	169
Abbildungsverzeichnis	176
Tabellenverzeichnis	178
Literaturverzeichnis	179
Danksagung	187
Kurzlebenslauf	188
Erklärung	189

1. Einleitung und Motivation

Geräuschimmissionen sind in der Regel stärker belästigend,
wenn sie *tonhaltig* sind.

Dieser einleitende Satz der DIN 45681:2006-08 „Akustik - Bestimmung der Tonhaltigkeit¹ von Geräuschen und Ermittlung eines Tonzuschlages für die Beurteilung von Geräuschimmissionen“ [THK06]² stellt die Motivation für die vorliegende Arbeit dar: Um die belästigende Wirkung von Geräuschimmissionen korrekt zu erfassen und zu bewerten, muss neben vielen weiteren Faktoren die Tonhaltigkeit (THK) berücksichtigt werden.

Die THK beschreibt das Hervortreten von Tönen oder anderen dominanten tonalen Schallkomponenten in einem zumeist rauschhaften Hintergrund. Da die THK z.B. von vielen mechanischen oder elektrischen Geräten und Maschinen hervorgerufen wird, stellt sie ein ernstes und nahezu alltägliches Problem in einer hochtechnisierten Industriegesellschaft dar.

Es ist daher ein wesentliches Ziel der DIN 45681 die Tonhaltigkeit zu objektivieren, um auf diese Weise Geräuschimmissionen adäquat messen und beurteilen zu können. Demgegenüber berücksichtigt die DIN 45681 das *Perzept* Tonhaltigkeit nur ansatzweise und die Erfassung der subjektiven *Wahrnehmung* der Tonhaltigkeit bleibt unklar.

Das große Interesse an der THK und ihre Wichtigkeit kann an der Anzahl der Studien zur THK abgelesen werden. So wurden während der letzten zehn Jahre mehr als 30 Untersuchungen allein auf der Deutschen Jahrestagung für Akustik (DAGA) präsentiert, die sich unmittelbar mit der Tonhaltigkeit beschäftigen.

Neben der Norm DIN 45681 selbst und ihrer allgemeinen Anwendung (siehe z.B. [Sag01], [PRA01] [Pio03] [SS03]) erforschen die eher anwendungsbezoge-

¹ Die Tonhaltigkeit wird im Verlauf dieser Arbeit als *THK* abgekürzt.

² Die DIN 45681:2006-08 ist die zweite Berichtigung der DIN 45681:2005-03 [THK05]. Diese beiden Versionen sind prinzipiell identisch, bis auf den Algorithmus zur automatischen Klassifizierung der THK. Im weiteren ist daher immer die neue Version der DIN 45681:2006-08 gemeint, wenn nicht explizit eine alternative Aussage getroffen wird.

nen Untersuchungen u.a. die Geräuschemissionen von Geräten mit elektrischen Motoren, Flug-, Schienen- und Straßenverkehr sowie Industrie- und Windkraftanlagen.

Darüber hinaus ist die THK ein wichtiger Faktor für das Sounddesign. Ein aktuelles Beispiel dazu ist die Gestaltung eines akustischen Signals für Elektroautos, welches dazu dienen soll, die Zustände des Fahrzeugs wie Geschwindigkeit, Beschleunigung, Fahrtrichtung usw. an die Umwelt zu übermitteln (<http://gm-volt.com/2010/06/11/nissan-leaf-electric-car-sounds-revealed/>).

Die THK ist auch Gegenstand von grundlagenorientierten Untersuchungen, die sich mit der Wahrnehmung und den zugrunde liegenden Eigenschaften der THK beschäftigen. Dazu gehören z.B. die Arbeiten von Hansen [Han10].

Die drei aufgeführten Bereiche zeigen den nach wie vor bestehenden Bedarf, die THK weiter zu erforschen und zu diskutieren. Auch die vorliegende Arbeit soll einen Beitrag dazu leisten.

Dafür wird im Folgenden zunächst kurz auf die aktuell gültige Fassung der DIN 45681:2006-08 eingegangen, um ihre Grundlagen zu erläutern und Schwächen zu thematisieren (Kapitel 2).

Im sich anschließenden Kapitel 3 werden perzeptuelle Aspekte der THK untersucht. Dazu wird zunächst ein geeignetes Verfahren identifiziert, um die Wahrnehmung der THK zu analysieren und den perzeptuellen Raum der Zuhörer abzubilden.

In Kapitel 4 wird ein geeigneter Orientierungsschall zur akustischen Demonstration der THK konzipiert und evaluiert, sowie eine systematische Probandeninstruktion für die subjektive Bewertungen der THK vorgeschlagen.

Die THK von Einzeltönen in Rauschen wird mithilfe vollständiger Paarvergleiche im Kapitel 5 untersucht, bevor ein wesentlich allgemeineres und flexibleres Verfahren zur Bewertung der THK entwickelt wird (Kapitel 6). Ein großer Vorteil des Verfahrens ist die direkte Interpretierbarkeit der gewonnenen Messergebnisse und die für die Probanden einfache Durchführbarkeit.

Das Verfahren wird evaluiert und die Parameter der THK damit ermittelt. In Kapitel 7 wird schließlich ein Verfahrenvergleich durchgeführt, um die verwendeten Methoden zu prüfen und zu optimieren. Dabei wird untersucht, wie die THK von Schallen mit mehreren tonalen Komponenten zusammengefasst werden kann und welche psychoakustischen Grundlagen eine Modellierung ermöglichen.

Abschließend werden die gewonnen Erkenntnisse kurz zusammengefasst (Kapitel 8).

2. DIN 45681

Im folgenden werden die für diese Arbeit relevanten Formeln und Gesetzmäßigkeiten der DIN 45681 kurz dargestellt.

Dabei wird nur an den ausdrücklich markierten Stellen zwischen der Version DIN 45681:2006-08 [THK05] und der 2. berichtigten Fassung DIN 45681:2006-08 [THK06] unterschieden, da die Abweichungen nur die Implementation des Algorithmus zur automatischen Klassifizierung der THK betreffen. Alle sonstigen Einzelheiten sind für beide Versionen identisch.

Alle nicht im Folgenden diskutierten Zusammenhänge und Gesetzmäßigkeiten haben keine Relevanz für die weitere Arbeit und müssen daher bei Bedarf in der Norm nachgeschlagen werden.

2.1. Messdurchführung

Soll die THK normgerecht anhand der DIN 45681 gemessen werden, etwa in der gutachterlichen Praxis, so sind die Vorgaben des angewandten Mess- und Beurteilungsverfahrens (z. B. der TA Lärm) hinsichtlich der Auswahl des Messortes, der Messzeit und -dauer, des Fremdgeräusches, usw. zu beachten.

In dieser Arbeit werden jedoch keine Messungen der THK durchgeführt, die spezielle Anforderungen diesbezüglich benötigen. Es werden keine weiteren normierten Messverfahren zusätzlich zur DIN 45681 verwendet.

2.2. Eingangswerte der THK-Berechnung

Messgröße für die Bestimmung der Tonhaltigkeit ist der Schalldruck $p(t)$, der mit einem Schallpegelmessgerät nach DIN EN 61672-1 mindestens der Klasse 2 aufgenommen wird. Die Pegelwerte $p(t)$ müssen A-bewertet vorliegen oder zu Beginn der Auswertung entsprechend umgerechnet werden.

Da keine gutachterlichen Messungen vorgenommen werden, wird aus den beiden folgenden Gründen im Weiteren komplett auf die A-Bewertung verzichtet: Einerseits wird die Bewertung sowohl für die Ton- wie auch für die Hintergrundpegel vorgenommen. Daher ist die Auswirkung auf die maßgebliche Differenz (siehe Kapitel 2.3) vernachlässigbar, ein zusätzlicher Erkenntnisgewinn wird für keine der beiden Methoden erwartet. Andererseits würde eine A-Bewertung die

Voraussetzungen der Experimente in den Kapiteln 6 und 7 beeinträchtigen, was möglichst vermieden werden soll.

Die Frequenzanalyse kann entweder mit analogen Geräten der entsprechenden Klasse oder digital mit geeigneter Computer-Hardware durchgeführt werden. In beiden Fällen muss der konstante Linienabstand Δf der Frequenzen im Spektrum zwischen 1,9 Hz und 4 Hz (einschließlich) liegen und das Hanning-Fenster zur Verringerung von Seitenbändern angewendet werden.

2.3. Maßgebliche Differenz ΔL

Der wichtigste Parameter bei der Bestimmung der THK ist die Differenz ΔL des Tonpegels L_T zum Pegel der Mithörschwelle $L_G + a_v(f)$. Die Mithörschwelle wird bestimmt durch das frequenzabhängige Verdeckungsmaß $a_v(f)$, siehe Anhang A.2, so dass gilt:

$$(2.1) \quad \Delta L = L_T - L_G - a_v(f).$$

Dabei ist L_T der Tonpegel in dB, L_G der Frequenzgruppenpegel des verdeckenden Geräusches (in dB) und $a_v(f)$ das frequenzabhängige Verdeckungsmaß (in dB). Der Verlauf des Verdeckungsmaßes wird durch Gleichung A.1 angegeben und ist in Abbildung A.2 über der Frequenz dargestellt. Abbildung A.2 ist zu entnehmen, dass $a_v(f)$ nur negative Werte annehmen kann.

Die Mithörschwelle³ kennzeichnet den Wert, bei dem ein einzelner Ton im Hintergrundgeräusch gerade an der Schwelle zwischen „wahrnehmbar“ und „nicht wahrnehmbar“ liegt. Die Differenz ΔL charakterisiert daher die Lautstärke oberhalb der Mithörschwelle.

In dieser Arbeit wird im weiteren die Größe ΔL vereinfachend und abkürzend nach DIN 45681:2005-03, Abschnitt 5.3.8, als „maßgebliche Differenz“ bezeichnet. Streng genommen muss zusätzlich beachtet werden, ob und wie die einzelnen maßgeblichen Differenzen gemittelt werden. Für die weitere Argumentation spielt das im Rahmen dieser Arbeit jedoch keine Rolle und kann daher vernachlässigt werden.

Die maßgebliche Differenz wird im Kapitel 5.5 dieser Arbeit untersucht.

2.4. Tonzuschlag K_T

Bei der Beurteilung von Geräuschimmissionen wird ein Tonzuschlag K_T vergeben, wenn die maßgebliche Differenz größer als Null ist: $\Delta L > 0$. Das bedeutet, dass es mindestens einen Ton gibt, der aus dem verdeckenden Geräusch hervortritt und somit wahrnehmbar ist.

³ Die Mithörschwelle wird synonym auch als Maskierungsschwelle oder Verdeckungsschwelle bezeichnet.

Die Höhe des Tonzuschlages K_T richtet sich nach der durch die Differenz ΔL gegebenen Deutlichkeit, mit der die tonale Komponente aus dem verdeckenden Geräusch hervortritt, siehe Tabelle 2.1.

Tabelle 2.1.: Tonzuschlag K_T in Abhängigkeit von ΔL

ΔL in dB	K_T in dB
$\Delta L \leq 0$	0
$0 < \Delta L \leq 2$	1
$2 < \Delta L \leq 4$	2
$4 < \Delta L \leq 6$	3
$6 < \Delta L \leq 9$	4
$9 < \Delta L \leq 12$	5
$12 < \Delta L$	6

Wie aus Tabelle 2.1 zu entnehmen ist, beträgt der Bereich der THK (also die maßgebliche Differenz) zwischen keinem und dem maximal möglichen Tonzuschlag K_T 12 dB. Für größere maßgebliche Differenzen wird der Tonzuschlag nicht mehr weiter erhöht.

2.5. Frequenzabhängigkeit der THK

In der aktuellen Version der DIN 45681:2005-03 wird die Frequenzabhängigkeit der THK ausschließlich durch das frequenzabhängige Verdeckungsmaß $a_v(f)$ berücksichtigt, vgl. Kapitel 2.3.

Ein Einfluss der Tonfrequenz über das Verdeckungsmaß hinaus existiert in der DIN 45681 nicht. Aus diesem Grund wird die Frequenzabhängigkeit der THK in Kapitel 5.6 und 6.4.8 analysiert, um so die Relevanz für die Wahrnehmung der THK zu evaluieren und die Annahme der Norm zu prüfen.

Weiterhin wird untersucht werden, ob eine untere Grenze für die Frequenz der Töne existiert, ab der sie keine THK mehr hervorrufen. Dies ist von Interesse, da für die gültige Version der DIN 45681:2005-03 die untere Grenze für die Tonfrequenz bei 90 Hz festgelegt ist.

2.6. Ausgeprägtheit von Bandpassrauschen

Die Ausgeprägtheit eines Tones hängt von der Bandbreite des Tones und seiner Flankensteilheit ab. Sind die entsprechenden Kriterien nicht erfüllt, so ist der Ton für Normalhörende nicht hörbar.

Die Wahrnehmung „Ton“ wird in der Norm auf Schmalbandrauschen erweitert. Es werden die gleichen Kriterien bzgl. der Ausgeprägtheit angenommen:

Bei einer Ausgeprägtheit eines auf Bandpassrauschen basierenden Tones von 70% relativ zur Ausgeprägtheit eines Sinustones, ergibt sich die maximal zulässige Bandbreite Δf_R in Abhängigkeit von der Tonfrequenz f_T angenähert zu:

$$(2.2) \quad \Delta f_R = 26 \left[1 + \frac{f_T}{1000 \text{ Hz}} \right] \text{ dB}$$

Diese Modellierung wird in Kapitel 6.4.5 überprüft. Da die Flankensteilheit nicht experimentell variiert wird, findet hier keine weitere Betrachtung der Flankensteilheit statt.

2.7. Genau zwei Töne in einer Frequenzgruppe

Treten in einer Frequenzgruppe genau zwei Töne mit den Tonfrequenzen f_{T_1} und f_{T_2} auf, so werden sie getrennt beurteilt, wenn jede Tonfrequenz unter 1000 Hz liegt und die Frequenzdifferenz f_D

$$(2.3) \quad f_D = |f_{T_1} - f_{T_2}| \quad f_{T_1}, f_{T_2} < 1000 \text{ Hz}$$

den folgenden Wert (für den Ton f_T mit der größeren maßgeblichen Differenz) überschreitet:

$$(2.4) \quad f_D = 21 \cdot 10^{1.2 \left| \log_{10} \left(\frac{f_T}{212 \text{ Hz}} \right) \right|^{1.8}} \text{ Hz} \quad 88 \text{ Hz} < f_T < 1000 \text{ Hz}$$

Dazu existiert in der DIN 45681:2005-03 die hilfreiche Abbildung D.1. Die angegebene Vorschrift wird experimentell in Kapitel 6.4.3 überprüft.

2.8. Mehr als zwei Töne in einer Frequenzgruppe

Mehr als zwei Töne in einer Frequenzgruppe werden im Folgenden nicht untersucht, daher wird das Vorgehen hier nicht weiter diskutiert.

2.9. Töne in unterschiedlichen Frequenzgruppen

Tonale Komponenten in unterschiedlichen Frequenzgruppen werden getrennt beurteilt. Zur Entscheidung, ob ein Tonzuschlag zu vergeben ist, wird in diesem Fall nur der am stärksten hervortretende Ton herangezogen. Treten innerhalb einer Frequenzgruppe mehrere Töne auf, so werden ihre Tonpegel L_{T-i} zu einem Tonpegel L_T energetisch addiert.

Ein Ton führt nur dann zu einem Tonzuschlag, wenn seine Ausgeprägtheit mindestens 70% beträgt. Dies bedeutet eine von der Tonfrequenz abhängige

maximale Bandbreite Δf_R (siehe Gleichung 2.3) und erfordert eine Flankensteilheit von mindestens 24 dB/Oktave.

Harmonische Vielfache eines Tones werden, unabhängig von diesem Ton, wie jede andere Komponente des Spektrums beurteilt.

Das Konzept, harmonische Vielfache wie jede andere Komponente des Spektrums zu behandeln und nur die am stärksten hervortretende Komponente zu berücksichtigen wird im Kapitel 6.4.1 und 7.5 untersucht.

2.10. Nichtstationäre Signale

Das Verfahren unterscheidet nicht zwischen stationären und nicht stationären Geräuschen. Die Norm ([THK06], Abschnitt 5.1, definiert daher: „Signale mit sehr hoher Pegel- und/oder Frequenzdynamik, die mit einer 3-Sekunden-Mittelung nicht mehr korrespondieren, können durch diese Norm nicht beurteilt werden.“

Das trifft auch für nicht eindeutig zu beschreibende tonale Komponenten zu. So sind für die DIN 45681 [THK06] „heulende“ oder „kreischende“ Signalanteile explizit von der algorithmischen Berechnung ausgenommen [THK06].

Signale mit derartigen Eigenschaften müssen daher weiterhin durch das subjektive Empfinden eines Gutachters bewertet und eventuelle Tonzuschläge festgelegt werden.

2.11. Probleme bei der Bestimmung des Schmalbandpegels

Für sehr große maßgebliche Differenzen ΔL ist der Algorithmus für die Trennung von Ton und Rauschen mit den in der Norm definierten Parametern wie z.B. die Frequenzauflösung Δf und insbesondere die Mindestzahl der Spektrallinien unter und über der untersuchten Tonfrequenz f_T nicht geeignet, vgl. Norm DIN 45681:2006-08 [THK06] Detaildiagramm 1 in Bild H.2.

Für die Anwendung der Norm im Rahmen ihres geplanten Einsatzbereiches ist dieser Umstand unproblematisch, da maßgebliche Differenzen in dieser Größe kaum mit der DIN 45681 beurteilt werden. Der Pegel der tonalen Komponente müsste so groß sein, dass andere Regularien (z.B. die TA Lärm) den Effekt dieses Tones adäquater beurteilen als die THK.

Für die Analyse der Stimuli in Kapitel E.2 verhindert dieser Sachverhalt allerdings eine sinnvolle Anwendung auf diejenigen Schalle mit großem ΔL .

3. Perzeptuelle Aspekte der Tonhaltigkeit

Die THK wird in der DIN 45681 implizit definiert, vgl. Kapitel 2. Auf diese Weise werden Vorschriften definiert und psychoakustische Gesetzmäßigkeiten verwendet, um die THK objektivieren und somit messen zu können.

Dieses Vorgehen ist in den Ingenieurwissenschaften weit verbreitet und wird im technisch-normativen Bereich häufig angewendet. Ähnliche Herangehensweisen werden beispielsweise auch für die psychoakustische Rauigkeit oder Schärfe benutzt.

Dabei wird die Wahrnehmung, die diesen (psychoakustischen) Größen zugrunde liegt, oftmals nicht ausreichend thematisiert und beachtet. Dies wird spätestens dann zum Manko, wenn die objektiv bestimmten Maße mit der menschlichen Wahrnehmung verglichen werden sollen.

Ein Grund für die Vernachlässigung der *wahrgenommenen* THK liegt darin, dass die THK kein *Begriff* des alltäglichen Lebens ist. THK ist eine genormte Beschreibung, zu der kein einheitlich verwendeter Begriff des allgemeinen Sprachgebrauchs korrespondiert. So ist das Wort „Tonhaltigkeit“ gar nicht im Duden [Dud00] verzeichnet, „tonhaltig“ wird ausschließlich als „Ton enthaltende Erde“ aufgeführt und nicht bzgl. seiner akustischen Bedeutung erwähnt.

Darüberhinaus ist der Begriff THK in keinem namhaften gedruckten Lexikon aufgeführt, was die Bedeutung der THK für den täglichen Gebrauch verdeutlicht: der Begriff ist offensichtlich nur für Spezialisten relevant. Diese haben zumeist eine derartige Ausbildung, dass sie die technische Definition der THK verstehen und interpretieren können.

Doch selbst internationale Spezialisten haben ein Problem mit der einheitlichen Begriffsbildung der THK in verschiedenen Sprachen. So existiert keine einheitliche Übersetzung des Begriffs THK ins Englische. Stattdessen werden verschiedene Umschreibungen benutzt, die einen tonalen Geräuschanteil vor einem (überwiegend rauschhaften) Hintergrund charakterisieren, vgl. dazu Hansen [Han10].

Vor diesem Hintergrund verwundert es nicht, dass Personen ohne entsprechende Hintergrundkenntnisse den Begriff THK gar nicht einordnen können oder sogar falsch interpretieren.

3. Perzeptuelle Aspekte der Tonhaltigkeit

Aus diesem Grund wird im Folgenden zunächst die komplexe Wahrnehmung⁴ der THK untersucht. Das führt zu dem allgemeinen Problem des Messens von Wahrnehmungen, bzw. des Messens der Ausgeprägtheit dieser Wahrnehmungen.

Abstrakte Wahrnehmungen, wie z.B. die THK, sind i.A. stark subjektiv und individuell geprägt und daher nicht direkt messbar. Die üblicherweise benutzten Messverfahren wie Fragebögen, semantische Differentiale, Skalierungen aller Art oder Interviews sind daher nur bedingt zur Messung der Wahrnehmungen geeignet.

Kruse [Kru05] untersucht diese Messverfahren von einem allgemeinen Standpunkt aus und thematisiert die möglichen Probleme des Messvorgangs an sich. Da seine Herangehensweise interessante Aspekte für die THK bietet, wird sein Standpunkt zunächst näher betrachtet.

Bei seiner Analyse thematisiert Kruse [Kru05] die folgenden Schwierigkeiten, die er zunächst als ausweglos ansieht und daher als Dilemmata bezeichnet:

1. Dilemma der *Irrationalität*

„Der Gegenstand der Messung entzieht sich (...) weitgehend dem bewussten Erleben und damit dem direkten Zugriff durch die Auskunft gebenden Personen.“

2. Dilemma der *Ambiguität*

„Die natürliche Mehrdeutigkeit und Kontextabhängigkeit sprachlicher Äußerungen macht die Interaktion zwischen Fragesteller und Auskunftsperson zu einem Akt gegenseitiger Interpretation.“

3. Dilemma der *Konstruktivität*

„Die Abhängigkeit des Gegenstandes der Messung von der individuellen Lerngeschichte der Auskunft gebenden Personen macht einen intersubjektiven Vergleich von Einzelaussagen und unabhängig voneinander erhobenen Merkmalsausprägungen nahezu unmöglich.“

Das Dilemma der Irrationalität (1) bedeutet bei der Aufgabenstellung, die wahrgenommene THK zu erfassen, dass die Testpersonen i.A. nicht wissen, welche Wahrnehmung mit dem für sie abstrakten Begriff THK beschrieben wird. In Skalierungs- und Bewertungsversuchen wissen sie daher nicht, was sie beurteilen sollen.

⁴ „Wahrnehmung bezeichnet im Allgemeinen den Vorgang der Sinneswahrnehmung von physikalischen Reizen aus der Außenwelt eines Lebewesens, also die bewusste und unbewusste Sammlung von Informationen eines Lebewesens über seine Sinne. Auch die so aufgenommenen und ausgewerteten Informationen werden Wahrnehmungen (oder Perzepte) genannt. Diese werden laufend mit den als Teil der inneren Vorstellungswelt gespeicherten Konstrukten oder Schemata abgeglichen“ [Wik09].

In einer unveröffentlichten Vorstudie konnte allerdings gezeigt werden, dass durch die Präsentation von akustischen Beispielen⁵ das Perzept THK verdeutlicht werden kann. Die Wahrnehmung des Phänomens, welches hier als THK bezeichnet wird, ist allen Testpersonen vertraut und verschiedene Ausprägungen der THK sind bekannt.

Bei der Erfassung der wahrgenommenen THK ergeben sich aus den genannten Dilemmata weitere Restriktionen für die Anwendung psychometrischer Messmethoden (nach Kruse [Kru05]): „Alle Befragungsformen, die mit der Vorgabe von ausformulierten Merkmalsbeschreibungen und Bewertungsdimensionen arbeiten, lassen das Dilemma der Ambiguität (2) und das Dilemma der Konstruktivität (3) ungelöst. Die Verfahren machen die eigentlich nicht zulässige Annahme, dass den befragten Personen eine Interpretation der vorgegebenen Beschreibungen und Dimensionen gelingt, die der Intention des Fragestellers entspricht. Die Aussagekraft der Ergebnisse ist abhängig von einem hochgradig unsicheren Akt der mehrfachen Übersetzung und Deutung (...). Darüber hinaus ignorieren alle Befragungsformen, die versuchen, Erwartungshaltungen, Einstellungen und Bewertungen von Menschen direkt zu erfassen, das Dilemma der Irrationalität (1). Die Verfahren machen die nur sehr eingeschränkt gültige Annahme, dass die befragten Personen selbst einen bewussten Zugang zu ihren entscheidungsrelevanten Präferenzen und Erwartungshaltungen besitzen. Die Tatsache, dass Auskunftspersonen dennoch scheinbar sicher Antwort geben, verleitet nur allzu schnell zu unzulässigen Verhaltensprognosen.“

Somit sind die häufig benutzten Standard-Verfahren wie z.B. das semantische Differential oder Fragebögen zur Erfassung der THK nicht ohne Weiteres geeignet, da sie die Deutungsabhängigkeit sprachlicher Kodierungen (Dilemma der Ambiguität) ignorieren. Das betrifft eine große Anzahl durchgeführter Studien.

Das Ziel muss daher zuallererst darin bestehen, an die individuellen und intuitiven Kriterien der Befragten heranzukommen, um ihre eigenen, zumeist unbewussten Kriterien offen zu legen. Dazu eignet sich z.B. ein offenes Interview, in dem es möglich ist, frei und individuell mit den Befragten zu interagieren.

Der Nachteil dieses sehr aufwendigen Verfahrens ist allerdings, dass es jeweils individuell ablaufen muss und sich daher in verschiedene Richtungen entwickeln kann. Somit ist eine Zusammenfassung der Daten für größere Gruppen nahezu ausgeschlossen [Kru05].

Zur Messung der Wahrnehmung THK wird daher eine Methode gesucht, die die inhaltliche Aussagekraft qualitativer Verfahren mit der Vergleichbarkeit und

⁵ Die Auswahl und Präsentation von akustischen Beispielen für das Perzept THK wird im nächsten Kapitel (4) ausführlich diskutiert.

Zusammenfassbarkeit standardisierter Fragebögen verbindet und somit die erläuterten Nachteile anhand der gestellten Anforderungen minimiert.

Eine Möglichkeit hierfür bietet das *Repertory-Grid* nach Kelly, welches auf seiner Theorie der persönlichen Konstrukte⁶ beruht, vergleiche z.B. [Kel55], [Sch92], [SC93], [Kru05] oder [MRB07]. Aufgrund der Inhaltsoffenheit des Verfahrens ist es geeignet, die persönlichen Begriffs- und Bedeutungsräume der Befragten aufzudecken, was ein wesentliches Ziel der nachfolgenden Untersuchung ist.

Ein bedeutender Vorteil dabei besteht darin, dass das Erhebungsritual eine absichtsvolle Ergebnisverzerrung nahezu ausschließt. Die komplexen Musterbildungen entstehen nur auf der Basis tatsächlich ausgebildeter Bewertungen und beziehen unbewusste Aspekte mit ein (Dilemma der Irrationalität) [MRB07].

3.1. Repertory-Grid Verfahren

Da das Repertory-Grid⁷ Verfahren für psychoakustische Untersuchungen eher ungewöhnlich ist, werden die relevanten Grundlagen für die Untersuchungen zur THK im Weiteren diskutiert.

3.1.1. Grundlagen

Persönliche Konstrukte beschreiben die Art und Weise eines Menschen, seine Welt wahrzunehmen und sich damit der Realität anzunähern. Sie dienen der Vorhersage und Kontrolle von Ereignissen. Da jeder Mensch im Laufe der Zeit sein individuelles Konstruktsystem aufbaut, wird die Reaktion auf ein und dasselbe Ereignis individuell unterschiedlich erlebt.

Die Konstrukte werden als „verbales Etikett“ für eine begriffliche Unterscheidung angesehen. Da wir üblicherweise in Kontrasten denken, ergibt der Begriff „gross“ keinen Sinn ohne eine Vorstellung von „klein“. Daher bestehen die Konstrukte aus (dichotomen) Kontrastpolen, die zusammengehören und gemeinsam analysiert werden müssen.

Das ist insbesondere wichtig, da je nach Kontext durchaus verschiedene Gegenpole zu einer verbalen Beschreibung existieren können. Diese unterschiedlichen Gegensatzpaare haben entsprechend verschiedene Bedeutungen [Kru05].

Zur Erfassung der individuellen Konstruktsysteme entwickelte Kelly die *Rep-*

⁶ Die Theorie der persönlichen Konstrukte besagt, dass es keine objektive Realität oder die absolute Wahrheit gibt, sondern jeder Mensch sich sein eigenes Weltbild konstruiert [SC93].

⁷ Das Repertory-Grid Verfahren wurde ursprünglich für die klinische Psychologie entwickelt. Aufgrund seines großen Potenzials hat sich der Anwendungsbereich in den letzten Jahren deutlich ausgedehnt, vor allem auf Forschungszwecke jenseits der Psychologie.

ertory-Grid Technik. Dabei werden dem Probanden mehrere Elemente⁸ präsentiert, für die er die Ähnlichkeiten und Kontraste zwischen ihnen benennen soll. Daraus werden im Weiteren Aussagen über die vom Probanden verwendeten Konstrukte⁹ und dem zugrunde liegenden Konstruktsystem ermittelt [SC93].

Eine große Stärke des Repertory-Grid Verfahrens liegt darin, die empirische Bedeutung der verwendeten Begriffe auf dem jeweils individuellen, d.h. subjektiven Bezugssystem zu erfassen und somit das „Alltagsverständnis“ der Befragten zu registrieren. Auf diese Weise ist es auch möglich, zunächst nicht sichtbare und fassbare Elemente und Tätigkeiten, d.h. implizites Wissen der Befragten, relativ leicht und strukturiert zu kategorisieren [MRB07].

Mit dem Repertory Grid ist es daher möglich, die Wahrnehmung der THK individuell und unbeeinflusst zu untersuchen, um die individuellen Wahrnehmungsräume der Teilnehmer identifizieren und vergleichen zu können.

3.1.2. Methode

Die Repertory-Grid Technik ist kein streng definierter Test im eigentlichen Sinne. Er stellt vielmehr ein Inventar von Verfahren bereit, die für jedes Problem eine andere, meist heuristische und kreative Nutzung des methodischen Prinzips ermöglicht [Sch92].

Das wesentliche Merkmal dieser Befragungstechnik besteht darin, dass die Teilnehmer zunächst jeweils selber einen individuellen, an die Befragungssituation angepassten „Fragebogen“ entwickeln, mit dem Sie im Folgenden ihre Bewertungen vornehmen.

Dabei stellen die Teilnehmer über ein sehr einfaches Befragungsritual Ereignisse oder Objekte ihrer Erfahrungswelt (die *Elemente*) auf der Grundlage von eigenständig beschriebenen Unterschieden oder Gemeinsamkeiten (genannt *Konstrukte*) in eine Abhängigkeitsbeziehung [Kru05].

Die Repertory-Grid Technik besteht im allgemeinen aus vier Schritten, vgl. Scheer [Sch92] und Kruse [Kru05]:

1. *Erhebungsphase*: Bestimmung der Elemente
2. *Evokationsphase*¹⁰: Einstiegsvergleiche und Benennung der polaren Konstruktdimensionen, d.h. Gewinnung der Konstrukte
3. *Beurteilungsphase*: Bewertung der Elemente bzgl. der Konstrukte
4. *Auswertung*

⁸ Die Elemente können z.B. Gegenstände, Personen, Ereignisse oder andere, auch abstrakte und unspezifische Dinge (wie im vorliegenden Fall akustische Stimuli) sein.

⁹ Die Konstrukte sind die Reaktionen, Verhaltensweisen, Bedeutungen oder wie im vorliegenden Fall die Wahrnehmungen der Elemente.

¹⁰ Evokation = allgemein das Erwecken bzw. Hervorrufen oder Wachrufen von Gedanken und Assoziationen.

3. Perzeptuelle Aspekte der Tonhaltigkeit

In diesem Zusammenhang wird das Repertory-Grid Verfahren benutzt, um die „verborgenen Begriffe“ bzgl. der Tonhaltigkeit aufzuklären. Dazu sollen die Teilnehmer akustische Stimuli, die verschiedene Aspekte bzgl. der Tonhaltigkeit abdecken, mit der zuvor beschriebenen Methode vergleichen.

Die Teilnehmer werden bewusst im Vorfeld der Untersuchung nicht über die THK aufgeklärt, da ihr eigenes, unverfälschtes Konstruktsystem bzgl. der THK offen gelegt werden soll. Dabei ist insbesondere die verbale Beschreibung der Unterschiede der Stimuli (d.h. indirekt der THK) von Interesse.

Im Hinblick auf die genannten Aspekte und Anforderungen werden die vier essentiellen Schritte des Repertory-Grid Verfahrens wie folgt konkret umgesetzt.

Erhebungsphase

In klassischen Untersuchungen mit dem Repertory-Grid müssen die Elemente, bzgl. derer das Konstruktsystem untersucht werden soll, zunächst durch die Teilnehmer individuell erarbeitet, bzw. festgelegt werden (Erhebungsphase).

Für die vorliegende Untersuchung entfällt dieser Schritt, da die Elemente (d.h. die zu untersuchenden, akustischen Stimuli) im Vorfeld der Untersuchung anhand der Fragestellungen und Hypothesen ohne Zutun der Teilnehmer definiert werden.

Die Aufgaben der Teilnehmer beginnen in der Untersuchung somit erst mit dem nächsten Schritt.

Evokationsphase

Jeweils drei der in der Erhebungsphase definierten Stimuli werden in randomisierter Reihenfolge zu sogenannten *Triaden* gruppiert. Sukzessive werden dann alle kombinatorisch möglichen Triaden nacheinander bearbeitet.

Die Aufgabe der Teilnehmer besteht darin, triadenweise zu entscheiden, welche beiden der drei dargebotenen Stimuli sich am ähnlichsten sind. Daraufhin müssen die Eigenschaften der Ähnlichkeit benannt werden. Schließlich muss die Eigenschaft benannt werden, durch die sich der verbleibende Stimulus von den beiden ähnlichen Stimuli unterscheidet.

Auf diese Weise wird jeweils mindestens ein Konstrukt bzgl. der Ähnlichkeit und ein Konstrukt bzgl. des Kontrasts der Stimuli verbal beschrieben und offen gelegt. Die so erhaltenen Konstrukte werden im Folgenden als Pole einer Skala mit der angegebenen Eigenschaft betrachtet.

In klassischen Repertory-Grid Verfahren müssen die Teilnehmer sich die Konstrukte teilweise erst selber erarbeiten, da sie nicht zwingend geläufig vorliegen. Während des Erarbeitens können verschiedene Effekte des Untersuchers auftreten. Einerseits kann er z.B. bei Nachfragen behilflich sein, Artikulationshilfen

geben oder triviale Konstrukte zurückweisen. Diese Möglichkeiten werden in der Regel die Datenqualität verbessern.

Andererseits kann der Untersucher die Teilnehmer auch (bewusst oder unbewusst) beeinflussen und so die Datenerhebung verfälschen. Um diesen negativen Effekt zu minimieren, werden die folgenden Untersuchungen ausschließlich scriptgesteuert per Computer, d.h. ohne menschlichen Untersucher, vorgenommen. Somit kann der Ablauf für alle Teilnehmer immer gleich und standardisiert durchgeführt werden. Die Oberfläche des in der Evokationsphase verwendeten Scripts zeigt Abbildung C.1 im Anhang C. An dieser Stelle sind auch die GUIs¹¹ für die weiteren Schritte des Repertory-Grid abgebildet und ihre Funktionalität wird beschrieben.

Beurteilungsphase

Die Bewertung der Elemente bzgl. der Konstrukte kann je nach Zweck der Untersuchung sehr unterschiedlich durchgeführt werden. In dieser Studie werden die Elemente auf zwei Arten anhand der Konstrukte skaliert bzw. bewertet:

1. Herstellung einer Rangordnung der Elemente bzgl. der Konstrukte
2. Kategoriale Bewertung der Elemente anhand der erhobenen Konstrukte.

Im ersten Fall besteht die Aufgabe der Teilnehmer darin, die Elemente, d.h. die akustischen Stimuli, nacheinander bzgl. jedes Konstruktes auf- bzw. absteigend anzuordnen, vgl. Abbildung C.2.

Im zweiten Fall wird die Ausprägung der Elemente auf einer X-stufigen Skala ermittelt. Die Pole der Skala sind mit den entsprechenden Konstruktpaaren, d.h. auf einer Seite das Konstrukt der Ähnlichkeit, auf der gegenüberliegenden das Konstrukt des Kontrastes bezeichnet.

Im konkreten Fall wird eine quasikontinuierliche Skala verwendet, d.h. die Anzahl der Stufen zwischen den Polen ist dem Teilnehmer nicht direkt ersichtlich, vgl. Abbildung C.3. Intern wird eine Abstufung von 0 bis 100 vorgenommen. Die Teilnehmer haben zu bewerten, welches der Konstrukte auf das zu bewertende Element zutrifft und welches Ausmaß es bzgl. dieses Konstruktes annimmt.

In den meisten Studien mit dem Repertory-Grid werden die Bewertungen idio-graphisch, d.h. genau auf diesen Teilnehmer zum Zeitpunkt der Untersuchung bezogen, ausgewertet. Für diese individuellen und sehr spezifischen Auswertungen wurde das Repertory-Grid Verfahren entwickelt und optimiert.

Abweichend davon werden in dieser Studie zwar die Konstrukte individuell erhoben, die Beurteilungen der Elemente aber anhand verallgemeinerter Konstrukte durchgeführt. Dieses Vorgehen ist dadurch motiviert, dass einerseits die individuelle „interne Darstellung“ der THK aufgedeckt, andererseits daraus aber eine objektive Beschreibung der THK abgeleitet werden soll.

¹¹ GUI = Graphical User Interface, d.h. die grafischen Benutzer-Oberflächen für die Teilnehmer.

3. Perzeptuelle Aspekte der Tonhaltigkeit

Die Durchführung des Repertory-Grid Verfahrens wird daher im konkreten Fall zweistufig vorgenommen. Nachdem die Konstrukte erhoben worden sind, erfolgt zunächst eine gemeinsame Auswertung über alle Teilnehmer hinweg. Erst danach werden die Elemente anhand der verallgemeinerten, d.h. über alle Probanden zusammengefassten Konstrukte, von allen Teilnehmern bewertet.

Zusätzlich werden die Elemente anhand von Konstrukten beurteilt, die nicht von den Teilnehmern selbst entwickelt worden sind, sondern zu Vergleichszwecken ergänzt werden.

Auswertung

Die Auswertung kann qualitativ oder quantitativ erfolgen und bzgl. der individuellen Daten oder der über die Teilnehmer zusammengefassten Daten erfolgen.

Zur qualitativen Auswertung werden vorrangig verschiedene inhaltsanalytische Verfahren wie Kontingenz- und Äquivalenzanalyse benutzt, um Beziehungen zwischen den gebildeten Kategorien (den Konstrukten) zu entdecken und zu untersuchen. Typische Fragestellungen bei dieser Art der Auswertung sind z.B.: „Welche Kategorien treten gemeinsam auf?“, „wie häufig treten bestimmte Kategorien auf?“ oder „welche Kategorien schließen sich gegenseitig aus?“.

Im Rahmen dieser Arbeit werden diese Fragestellungen nur insoweit verfolgt, wie sie Aussagen über die „interne Repräsentation“ der THK ermöglichen und das Vokabular zur subjektiven Beschreibung der THK offen legen. Darüber hinaus gehende Analysen werden nicht durchgeführt.

Die meisten Konstrukte haben zumindest eine allgemein übliche, „normale“ Bedeutung. Diese als Denotation¹² bezeichnete Bedeutung des Konstruktes wird von vielen Leuten geteilt und ist die gebräuchlichste Bedeutung des Konstruktes.

Konstrukte können aber auch verschiedene „subjektive“ Bedeutungen besitzen, die ihre Entstehung in den ganz persönlichen Erfahrungen jedes einzelnen Menschen haben. Diese Bedeutungen des Konstruktes werden Konnotationen¹³ genannt.

Obwohl jeder Einzelne viele Konstrukte mit persönlichen, konnotativen Bedeutungen belegt, gibt es für die meisten Konstrukte eine allgemein geteilte Bedeutung. Diese Idee von einer gemeinsamen, denotativen Bedeutung ist einer der Schlüssel für sinnvolle menschliche Kommunikation. [Med10]

Dieser Gedanke der gemeinsamen, denotativen Bedeutung stellt eine wichtige Grundlage der folgenden Auswertungen dar. Unter der Annahme, dass das Perzept THK von allen Teilnehmern ähnlich empfunden wird, soll die verbale

¹² Duden [Dud00]: sprachwissenschaftlich, die begriffliche oder Sachbedeutung eines Wortes.

¹³ Duden [Dud00]: sprachwissenschaftlich, die mit einem Wort verbundene zusätzliche Vorstellung.

Beschreibung der internen Repräsentation des akustischen Reizes THK individuell erfasst werden, um sie anschließend verallgemeinern zu können.

Repertory Grids können prinzipiell in vielfältiger Weise quantitativ analysiert werden. Entscheidend für den Erfolg der Auswertung ist die Möglichkeit, die Ergebnisse sinnvoll anhand der untersuchten Objekte zu interpretieren. Die Wahl des Verfahrens für die Auswertung muss daher immer im speziellen Einzelfall erfolgen.

Grundsätzlich sind folgende Verfahren geeignet, die Daten des Repertory Grid zu analysieren:

- Die (nichtparametrische) Faktorenanalyse, mit der die Komplexität des Datensatzes reduziert werden kann und speziell die Hauptkomponentenanalyse. Ziel dabei ist es, die zugrunde liegenden Dimensionen des Konstruktsystems des Teilnehmers (individuell) oder aller Teilnehmer (zusammengefasst) zu erkennen und darzustellen.
- Die Clusteranalyse, die typische Muster der Elemente erfasst.
- Die multidimensionale Skalierung (MDS), die Distanzen zwischen den Elementen berechnet und so einen mehrdimensionalen Raum für die Elemente definiert. Dieser Raum wird im Folgenden als „Wahrnehmungsraum“ (der THK) bezeichnet.
- Korrelationen können Beziehungen zwischen den Konstrukten bzw. den Elementen aufzeigen. Das kann z.B. im Hinblick auf Inkonsistenzen interessant sein.

Mithilfe der aufgezählten Verfahren kann eine „Profilbildung“ erfolgen, die für viele Untersuchungen wertvolle Ergebnisse liefert. Eine vollständige Auswertung aller Verfahren würde im Rahmen dieser Studie zu weit gehen. Da an dieser Stelle vorrangiges Interesse am Wahrnehmungsraum der THK herrscht, wird die MDS mit höchster Priorität ausgewertet.

Das vorliegende Set von Ähnlichkeitsmassen zwischen den Elementen (d.h. den Stimuli) wird dabei verwendet, die zugrunde liegende Struktur der Elemente zu ermitteln. Dies wird durch das Zuweisen der Elemente zu bestimmten Positionen im Wahrnehmungsraum¹⁴ der THK erreicht. Im Folgenden wird dafür das Verfahren PROXSCAL [SPS09] nach Commandeur und Heiser [CH93] in der SPSS-Implementation [SPS02] verwendet.

Der Wahrnehmungsraum wird zunächst individuell für sämtliche Teilnehmer ermittelt. Anschließend werden die individuellen Wahrnehmungsräume solange aufeinander rotiert, bis ein minimaler Abstand aller Elemente über die Wahrnehmungsräume hergestellt ist. Auf diese wird der gemeinsame Wahrnehmungsraum für alle Teilnehmer identifiziert.

¹⁴ Als Wahrnehmungsraum wird hier ein konzeptueller Raum mit möglichst wenigen Dimensionen bezeichnet, in dem die Abstände zwischen den Punkten des Raumes mit den vorgegebenen Ähnlichkeiten so gut wie möglich übereinstimmen.

Bei Bedarf können nun die zu den Elementen korrespondierenden Wortverwendungen im gemeinsamem Wahrnehmungsraum neu interpretiert werden. Die inhaltsanalytische Interpretation wird dadurch mathematisch unterstützt.

Es bleibt zusammenfassend festzuhalten, dass durch Verwendung des Repertory-Grid Verfahrens die von Kruse [Kru05] thematisierten und eingangs des Kapitels diskutierten Probleme umgangen werden können.

Da die Teilnehmer die Konstrukte zur Beschreibung der Elemente (nur) durch ihre eigenen individuellen Erfahrungen und ihren Kenntnissen entsprechend unter Verwendung ihrer spezifischen Ausdrucksweise und Wortwahl erzeugen, kann auf diese Weise das Dilemma der Irrationalität vermieden werden.

Andererseits erlaubt die Analyse der Ähnlichkeit der individuellen Beschreibungen, bzw. der individuellen Wahrnehmungsräume, trotzdem ein Verständnis der Bedeutung über die Teilnehmer hinweg (Lösung für das Dilemma der Ambiguität und das Dilemma der Konstruktivität). „Auf diesem Wege werden Tiefenstrukturen in den Präferenzen und Erwartungshaltungen von Gruppen zugänglich, die sonst über die Vorgabe von Beschreibungsdimensionen vom Rauschen der Interpretationsproblematik verdeckt werden“ [Kru05].

3.2. Setup der Messungen

3.2.1. Probanden

An den folgenden Untersuchungen nehmen insgesamt acht weibliche und acht männliche Probanden teil, die keinerlei Vorerfahrung mit psychoakustischen Experimenten haben. Im Laufe der Studie wird zudem sicher gestellt, dass sie den Begriff THK im Vorfeld der Messungen nicht kennen.

Das Alter der Probanden liegt zwischen 23 und 65 Jahren (Median 34 Jahre), sie nehmen freiwillig an den Untersuchungen teil und werden für Ihren Aufwand nicht finanziell entschädigt.

Nicht alle Probanden führen alle folgenden Bewertungen aus. Sie werden vielmehr gleichmäßig auf die einzelnen Untersuchungen aufgeteilt.

3.2.2. Durchführung

Die Stimuli (d.h. die Schallsignale) sind jeweils zwei Sekunden lang, ihre Darbietung erfolgt mittels eines elektrostatischen Kopfhörers (Sennheiser HE 60 / HEV 70), der direkt an den Kopfhörer-Ausgang eines Laptops angeschlossen ist.

Die Lautheit der Stimuli beträgt für alle Versuche 25 sone_{GF}¹⁵. Zur Einstellung

¹⁵ Die Einheit sone_{GF} bezeichnet die für Freifeldbedingungen („F“) und auf Grundlage von Frequenzgruppen nach Zwicker („G“) berechnete Lautheit [Hea10].

der Stimuli auf diese Lautheit wird das Head Acoustics Analysesystem Artemis [Hea98] verwendet.

Aufgrund der recht großen Lautheit¹⁶ ist es nicht erforderlich, die Untersuchungen in einem speziell akustisch abgeschirmten Raum durchzuführen. Stattdessen werden die nachfolgenden Skalierungen bei den Probanden zu Hause, in einem ruhigen und ungestörten Raum, vorgenommen.

Menzel [MRB07] empfiehlt die Verwendung von mindestens sechs Elementen für die sinnvolle Durchführung des RepGrid-Verfahrens. Sollen alle möglichen Triaden für N Elemente vollständig gebildet und bewertet werden, so steigt der Aufwand während der Evokationsphase proportional zu N^3 an¹⁷. Als Kompromiss zwischen Aufwand und Nutzen werden daher im Folgenden $N=6$ Elemente verwendet und alle möglichen Triaden berücksichtigt. In diesem Fall ergeben sich auf diese Weise 20 Triaden.

Die Probanden haben die Aufgabe, den Repertory-Grid Versuch mittels der im Anhang C präsentierten und diskutierten Benutzeroberflächen durchzuführen. Dafür wird ihnen das Verfahren erläutert, nicht allerdings das Ziel der Untersuchung. Ebenso werden sie über die Stimuli im Unklaren gelassen, um sie im Vorfeld der Studie nicht zu beeinflussen.

Vor den eigentlichen Untersuchungen muss jeder Proband eine Übung absolvieren. Das Verfahren an sich wird auf diese Weise demonstriert und trainiert, die Daten werden jedoch nicht aufgezeichnet. Dazu werden reine Sinustöne unterschiedlicher Frequenz bei einer für die Probanden angenehmen Lautstärke als Stimuli verwendet. Die Frequenzen der sechs Sinustöne betragen 200, 300, 400, 500, 600 und 700 Hz.

Der Übungsdurchgang wird nur so lange durchgeführt, bis die Probanden verstanden haben wie die Messung abläuft. Danach wird die Übung abgebrochen und die folgenden Messungen durchgeführt.

3.3. Untersuchung der maßgeblichen Differenz

Nachfolgend wird die maßgebliche Differenz ΔL untersucht, da sie der wichtigste Parameter der THK ist und somit größten Einfluss auf die THK besitzt.

3.3.1. Stimuli

Die Stimuli werden so ausgewählt, dass sie den maximal möglichen Bereich der maßgeblichen Differenz abdecken, siehe Tabelle 3.1.

¹⁶ Die A-bewerteten Schalldruckpegel der Stimuli liegen zwischen 68 dB(A) und 87 dB(A), siehe dazu die Übersicht der in den jeweiligen Untersuchungen verwendeten Stimuli.

¹⁷ Die Anzahl der Triaden wird berechnet als $\frac{N(N-1)(N-2)}{6}$, das entspricht der Pyramidenzahl $N - 2$

3. Perzeptuelle Aspekte der Tonhaltigkeit

Tabelle 3.1.: Stimuli für die Untersuchung der maßgeblichen Differenz mit dem Repertory Grid. N bezeichnet die Lautheit der Stimuli, der korrespondierende Pegel ist sowohl ungewichtet L_L als auch A-bewertet L_A angegeben. ΔL_t bzw. ΔL_g enthält die theoretische, bzw. die durch die DIN 45681:2006-08 berechnete maßgebliche Differenz der Stimuli und K_T den entsprechenden THK-Zuschlag laut Norm. Durch Anklicken der Beschreibung wird bei geeigneter Konfiguration und verfügbarem Zugriff auf das Internet der jeweilige Stimulus abgespielt. Alternativ können die Stimuli auch direkt von der Webseite www.tonhaltigkeit.de herunter geladen oder angehört werden.

Nr.	Beschreibung	N in sone _{GF}	L_L in dB	L_A in dB (A)	ΔL_t in dB	ΔL_g in dB	K_T in dB
1	nur Ton	25,1	88,4	87,0	∞	75,5	6
2	Ton + GAR	25,0	81,1	79,7	~ 40	36,7	6
3	Ton + GAR	25,0	70,1	69,1	13	10,2	5
4	Ton + GAR	25,0	69,8	68,8	6	3,6	2
5	Ton + GAR	25,0	69,7	68,8	1	—	0
6	nur GAR	25,0	69,8	68,8	0	—	0

Um die maßgebliche Differenz zu variieren, wird ein Ton in einem Hintergrundrauschen mit verschiedenen Pegeln addiert. Es ergeben sich so die in der Tabelle aufgeführten, theoretisch zu erwartenden maßgeblichen Differenzen ΔL_t .

Der verwendete Ton ist ein (reiner) Sinuston mit einer Frequenz von 700 Hz¹⁸, das Rauschen ist gleichmäßig anregendes Rauschen (GAR)¹⁹, siehe die Beschreibung im Anhang A.

Auf diese Weise ergibt sich ein Kontinuum der THK, d.h. der maßgeblichen Differenz. Stimuli Nr. 6 ist minimal tonhaltig, bzw. theoretisch gar nicht tonhaltig, da er ausschließlich aus Rauschen besteht. Für Stimuli Nr. 5 ist der Ton bei genauem Hinhören schwach, für Stimuli Nr. 4 problemlos wahrnehmbar. Für Stimulus Nr. 3 ist der Ton im Vergleich zum Hintergrund nochmal deutlicher als für die zuvor genannten Stimuli. Der Stimuli 5 besteht im Wesentlichen aus dem reinen Ton, das Rauschen hat einen vergleichsweise geringen Pegel. Der Stimuli mit der ausgeprägtesten maßgeblichen Differenz (Nr. 1) enthält ausschließlich den Ton. Seine theoretische maßgebliche Differenz ΔL_t ist mit ∞ angegeben, um darauf hinzuweisen, dass er die größte praktisch zu erzeugende maßgebliche

¹⁸ Die Motivation genau dieser Frequenz wird in Kapitel 6.2.2 diskutiert.

¹⁹ Die Motivation für die Wahl dieses Rauschens wird ausführlich in Kapitel 6.2.2 erläutert.

Differenz besitzt²⁰.

Da an dieser Stelle breitbandige Stimuli (prototypisch Nr. 6) mit schmalbandigen (Nr. 1) verglichen werden, sind die Signale gleich laut, obwohl ihre Pegel sich um bis zu 18 dB unterscheiden. Das ist in Einklang mit Daten von Zwicker ([Zwi82], Abb. 5.6).

3.3.2. Evokationsphase: Auswertung der Konstrukte

Die zehn Teilnehmer bewerten jeweils alle möglichen Triple, die aus den sechs Stimuli zusammengestellt werden können. Somit sollten sie im Normalfall 20 Konstrukte für Gemeinsamkeiten und 20 Konstrukte für die Kontraste der sechs Elemente angeben.

Die so gewonnenen Konstrukte werden sinnvoll zusammengefasst und klassifiziert (kodiert), indem die individuell verwendeten Begriffe der Probanden zunächst auf ihre Bedeutung (im Sinne von Synonymen) zurückgeführt werden. Daraufhin wird die Häufigkeit der jeweiligen Bedeutung ermittelt. Eine Auswertung, die primär auf die individuellen Daten ausgerichtet ist, erfolgt hier hingegen nicht.

Die zehn Probanden haben unterschiedliche Synonyme für den Begriff *Ton* gewählt: sechs Probanden benutzen „Ton“, drei weitere „Piepsen“ bzw. „Piepton“ und einer „Pfeifen“. Diese verschiedenen Begriffe werden im Weiteren unter dem Oberbegriff „Ton“ zusammengefasst.

Der Begriff Ton wird zusätzlich mit verschiedenen Adjektiven versehen wie z.B. „heller Ton“, „hoher Ton“ oder „schriller Ton“. Teilweise werden die Adjektive auch vergleichend gebraucht, indem ihre komparative Form verwendet wird.

Für das Rauschen haben alle zehn Probanden einheitlich den Begriff „Rauschen“ benutzt und typischerweise mit folgenden Adjektiven versehen: „gleichmäßiges Rauschen“, „dumpfes Rauschen“, „lautes Rauschen“. Einige Probanden haben die Rauschen mit Assoziationen belegt wie z.B. „Rauschen wie Wasser“, „Wasserfall“ usw.

Die Auswertung erfolgt im weitesten Sinne entsprechend der „Grounded Theory“ [Wik10]. Der Datensatz wird sukzessive kodiert und nach jeder Kodierung einerseits der Effekt und andererseits die Relevanz der Kodierung geprüft. Ist die Kodierung zielführend, wird sie beibehalten, andernfalls verworfen und durch eine alternative Kodierung ersetzt. Auf diese Weise wird durch einen mehrstufigen Prozess die Kodierung nach und nach vervollständigt und perfektioniert.

Bei Verwendung der beiden Synonyme „Ton“ und „Rauschen“ ergeben sich durch Auszählen der gewonnenen Konstrukte die in Tabelle 3.2 aufgelisteten

²⁰ Da die „echte“ maßgebliche Differenz erheblich von der verwendeten Hardware und deren Konfiguration abhängt, wird die theoretisch maßgebliche Differenz an dieser Stelle nicht genauer spezifiziert und symbolisch als ∞ bezeichnet.

3. Perzeptuelle Aspekte der Tonhaltigkeit

Konstruktpaare (Antonyme). In der dritten Spalte der Tabelle ist der relative Anteil angegeben, mit der diese Paare für die Generierung der Konstrukte verwendet wird.

Tabelle 3.2.: Während der Evokationsphase gefundene Konstrukte, hier als Kontrastpaare gegenübergestellt. In der rechten Spalte ist die relative Häufigkeit der Kontrastpaare aufgeführt. ☒ bezeichnet implizite Konstruktpaare, * Adjektive, die auch komparativ benutzt werden.

Gruppe	Konstrukt- bzw. Kontrastpaar	Anteil
Gruppe 1 Σ=37,5% „Extreme“	Ton – Rauschen	26,5%
	Ton – kein Ton	5,0%
	kein Rauschen – Rauschen	4,5%
	nur Ton – nur Rauschen	1,5%
Gruppe 2 Σ=38% „Mischung“	Rauschen+Ton – Rauschen	23,5%
	Rauschen+Ton – nur Ton	5,0%
	Ton – nur Rauschen	4,0%
	Rauschen – nur Ton	3,5%
	Rauschen+Ton – kein Rauschen zusätzlicher Ton☒	1,5% 0,5%
Gruppe 3 Σ=22,5% „Ausprägung“	starker* Ton – schwacher* Ton	5,0%
	lauter* Ton – leiser* Ton	3,5%
	Rauschen im Vordergrund – Hintergrund	3,5%
	Ton im Vordergrund – Hintergrund	3,0%
	Ton dominant – Rauschen dominant	2,0%
	Ton überwiegt☒	2,0%
	Ton klarer – unklarer	1,5%
	mehr Rauschen – weniger Rauschen	0,5%
	zu viel Rauschen – zu viel Ton	0,5%
	scharfer Ton – unscharfer Ton	0,5%
klarer Ton – unscharfer Ton	0,5%	
„Unbekannt“	Andere	2,0%

An einigen Stellen werden implizite Konstruktpaare benutzt. Diese sind durch ☒ markiert. So bedeutet z.B. „zusätzlicher Ton“ indirekt, dass dieser Ton beim Gegenpol des Konstruktes nicht vorhanden ist. Die mit * bezeichneten Adjektive werden als Komparativ benutzt, um auf diese Weise die Kontraste auszudrücken.

Im Wesentlichen können die auf Synonyme zurück geführten Konstruktpaare in drei Gruppen zusammengefasst werden:

Gruppe 1: Es wird unterschieden, ob ein Merkmal (Ton bzw. Rauschen) vorhanden ist oder nicht, bzw. in welche der beiden Gruppen (Ton bzw. Rau-

schen) das Signal einzuordnen ist.

Die Klassifizierung der Elemente ist quasi „binär“: entweder das Merkmal ist vorhanden oder nicht. Da diese Klassifizierung nur für die beiden extremen Stimuli 1 und 6, bzw. mit leichten Einschränkungen auch für Stimuli 2 und 5 möglich ist, wird diese Gruppe als „Extreme“ bezeichnet, um diesen Sachverhalt zu verdeutlichen.

Gruppe 2: Diese Gruppe stellt eine Erweiterung der zuvor genannten Gruppe dar. Zu den beiden extremen Kategorien „Ton“ und „Rauschen“ wird zusätzlich die „Mischung“ dieser beiden Kategorien eingeführt und als neue Kategorie „Rauschen+Ton“ verwendet.

Gruppe 3: Die dritte Gruppe ist erneut eine Erweiterung der vorigen Gruppe. Sie enthält die Konstrukte, bei der die Ausprägung des Tones unterschieden wird, in selteneren Fällen auch die Ausprägung des Rauschens.

Sechs der zehn Probanden beziehen ihre Urteile überwiegend auf den Ton, die anderen vier bilden ihre Konstrukte mit gleicher Häufigkeit auf das Rauschen.

3.3.3. Evokationsphase: Auswertung durch MDS

Bei einem vollständigen Vergleich aller Tripple erhält man durch die Benennung der beiden ähnlichen Stimuli eine vollständige *Ähnlichkeitsmatrix*, sowie einen *Kontrastvektor* aufgrund des verbleibenden Stimulus.

Mit diesen Daten können anhand der multidimensionalen Skalierung (MDS) Aussagen über die Abstände der Elemente (Stimuli) gemacht werden, vgl. z.B. [BEPW00], [Kru05] oder [SC93].

Auf diese Weise wird der individuelle „Wahrnehmungsraum“ der THK bzgl. der Stimuli abgebildet. Als ein spezielles Verfahren der MDS wird PROXSCAL (Proximity Scaling, vgl. Commandeur und Heiser [CH93] und [SPS09]) in der SPSS-Implementation (SPSS 11.5 [SPS02]) angewendet.

Dafür wird zunächst die Ähnlichkeitsmatrix in eine Distanzmatrix umgerechnet, indem alle Vorzeichen der Elemente der Ähnlichkeitsmatrix invertiert werden. Auf diese Weise werden ähnlich bewerteten Stimuli kleine und unähnlich bewerteten Stimuli große Distanzen zugeordnet.

Die weitere Auswertung mit PROXSCAL kann entweder auf ordinalem oder auf Intervallskalenniveau erfolgen. Da beide Methoden annähernd die gleichen Ergebnisse liefern, wird im weiteren Intervallskalenniveau angenommen. Das ist angesichts der Verwendung von Abständen der Distanzmatrix gerechtfertigt.

Um eine ungefähre Abschätzung der Dimensionalität des Wahrnehmungsraumes zu erhalten, wird der Scree-Test (auch „Ellenbogenkriterium“ genannt) nach Cattell [Cat66] verwendet. Das Verfahren ist nicht unumstritten ([Bor93]), aber

3. Perzeptuelle Aspekte der Tonhaltigkeit

verhältnismäßig einfach anzuwenden und liefert für die folgenden Analysen befriedigende Ergebnisse.

Beim Scree-Test wird der Stress²¹, d.h. ein Maß für die Varianzaufklärung über die Dimensionen, beginnend mit einer einzigen Dimension, aufgetragen. In Abbildung 3.3 ist exemplarisch der Screeplot für die vorliegenden Daten, zusammengefasst für alle Probanden, dargestellt.

Es wird nur die Anzahl der Dimensionen berücksichtigt, die bis zum „Ellenbogen“, d.h. dem Knick des Screeplots, reichen. In Abbildung 3.3 sind das genau zwei Dimensionen.

Für die so ermittelte Dimensionalität wird mittels des in PROXSCAL enthaltenen Simplex-Algorithmus ([SPS02]) der Stress minimiert und damit der gemeinsame Raum, d.h. die Distanzen zwischen den Stimuli, berechnet [CH93].

Als Gütemaß der Anpassung bzgl. der gewählten Dimensionalität können abschließend verschiedene Stress-Maße herangezogen werden. So berücksichtigen Stress-I, Stress-II und S-Stress die mangelnde Anpassung der Daten, während die berücksichtigte Streuung (DAF = Dispersion accounted for) und der Kongruenzkoeffizient nach Tucker die Anpassung messen. Niedrigere Stress-Maße (bis zum Minimum 0) und höhere Anpassungsmaße (bis zum Maximum 1) weisen auf bessere Lösungen hin ([CH93] und [BEPW00]).

In Abbildung 3.1 sind die so berechneten, individuellen Wahrnehmungsräume für die zehn Probanden, jeweils mit den Buchstaben A–J bezeichnet, gemeinsam in eine Abbildung eingetragen²². Die Zahlen neben den Buchstaben bezeichnen jeweils die Stimuli. Zur besseren Übersicht sind die Stimuli probandenweise in aufsteigender Reihenfolge mit Linienzügen verbunden.

Die Dimensionalität der individuellen Wahrnehmungsräume wird durch visuelle Inspektion der Scree-Plots ermittelt. Für neun der zehn Probanden sind die Wahrnehmungsräume zweidimensional, für den Proband „H“ dagegen nur eindimensional. Die Urteile des Probanden „H“ bewegen sich daher lediglich entlang der Dimension 1.

Das Repertory-Grid bietet die Möglichkeit, die verwendeten Konstruktpaare im Wahrnehmungsraum anzuordnen und für bestimmte Bereiche des Wahrnehmungsraumes zu vergleichen. Diese Analyse wird hier allerdings nicht weiter verfolgt.

Da die Probanden offensichtlich eine sehr ähnliche Repräsentation ihres Wahrnehmungsraumes aufweisen, ist davon auszugehen, dass die zugrunde liegenden Bewertungen ebenfalls sehr ähnlich sein müssen. Auf dieser Basis wird der gemeinsame Wahrnehmungsraum für alle Probanden durch Mittelung der Ähnlichkeitsmatrizen erzeugt, siehe Abbildung 3.2.

²¹ Stress ist das Akronym für STandardized REsidual Sum of Squares.

²² Dafür ist der Wahrnehmungsraum so „gespiegelt“, dass Stimulus 1 immer unten rechts in der Abbildung liegt.

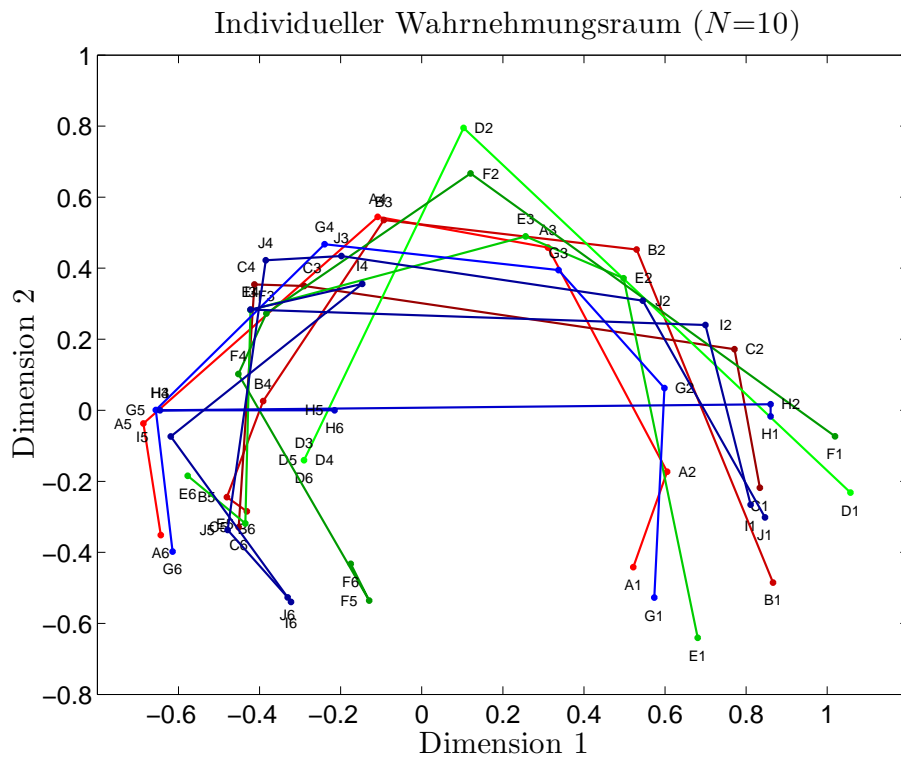


Abbildung 3.1.: Individuelle Wahrnehmungsräume der Versuchspersonen für die THK. Die Buchstaben kodieren die zehn Probanden, die Zahlen die sechs Stimuli wie in Tabelle 3.1 definiert.

Die erklärte Streuung (D.A.F.) für diesen gemeinsamen, zweidimensionalen Raum beträgt .999. Das zeigt, dass weitere Dimensionen den Wahrnehmungsraum nicht besser abbilden können. Das gleiche Ergebnis zeigt auch der Scree-Plot für den gemeinsam Wahrnehmungsraum, Abbildung 3.3. Da die individuellen Wahrnehmungsräume zweidimensional sind, sollte das auch für den gemeinsamen Wahrnehmungsraum gelten.

Abbildung 3.2 zeigt im wesentlichen drei Cluster: je einer für die Stimuli 1 und 2, bei denen der Ton deutlich dominiert, Stimuli 3 und 4, die eine Mischung aus Ton und Rauschen mit mittlerer THK aufweisen, sowie Stimuli 5 und 6, für die das Rauschen überwiegt, bzw. die keinen Ton enthalten.

Die durch Auswertung der Konstrukte ermittelten Gruppen, vgl. Tabelle 3.2, finden sich auch in Abbildung 3.2 wieder. Hier entspricht der Gruppe 1 „Extreme“ die Dimension 1 des Wahrnehmungsraumes, die Stimuli sind von „Rauschen“ auf der linken Seite zu „Ton“ auf der rechten Seite des Wahrnehmungsraumes angeordnet. Gruppe 2 („Mischung“) entspricht im Wahrnehmungsraum der Dimension 2. Die „reinen“ Stimuli („nur Ton“ bzw. „nur Rauschen“) befinden sich im unteren Bereich, die „gemischten“ Stimuli im oberen Bereich des

3. Perzeptuelle Aspekte der Tonhaltigkeit

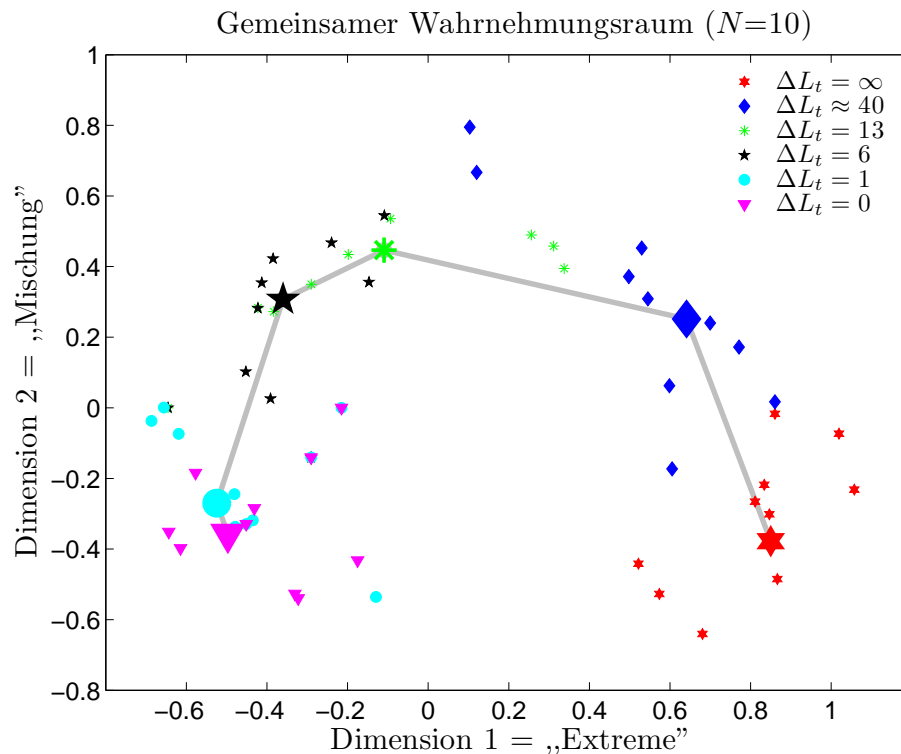


Abbildung 3.2.: Über die zehn Teilnehmer gemittelter Wahrnehmungsraum für die THK. Die Punkte der individuellen Wahrnehmungsräume sind aus Abbildung 3.1 übernommen.

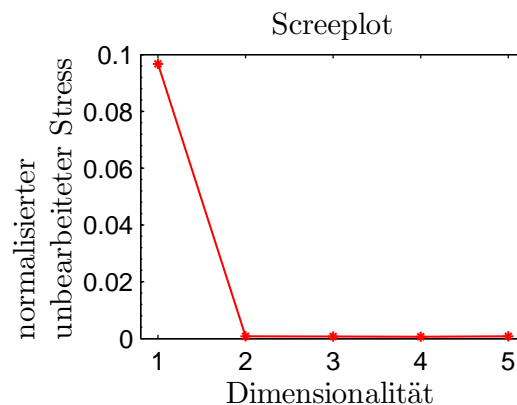


Abbildung 3.3.: Scree-Plot für den gemittelten Wahrnehmungsraum

Wahrnehmungsraumes. Die „Ausprägung“ (Gruppe 3) wird vorrangig benutzt, um die Unterschiede der Stimuli 3, 4 und 5, also der „gemischten“ Stimuli zu beschreiben. Die technische THK fällt entlang der grau eingezeichneten Linie von Stimulus 1 auf der linken Seite zu Stimulus 6 kontinuierlich ab. Dabei ha-

ben die Probanden für den mittleren Bereich der THK (Stimuli 3–5) eine in etwa eindimensionale interne Repräsentation der Stimuli. Für das Maximum und Minimum der THK ist eine eindimensionale Beschreibung der THK im Wahrnehmungsraum dagegen nicht mehr möglich.

3.3.4. Rangordnung der Elemente (Stimuli)

Der nächste Teil des Repertory-Grids besteht aus einer Rangordnung, d.h. Anordnung der Elemente entlang der gefundenen Skalen (i.e. die Kontrastpaare). Für die meisten Untersuchungsziele ist es erforderlich, die Rangordnung individuell anhand der spezifischen Konstruktpaare durchzuführen.

Stattdessen wird die Skalierung hier nur für die zusammengefassten Konstrukte durchgeführt. Dafür wird zum einen die Skala mit dem Kontrastpaar „Ton – Rauschen“ als Prototyp der Gruppe „Extreme“ verwendet, vgl. Tabelle 3.2. Weiterhin wird die Rangordnung bezüglich der Skala „Ton mit Rauschen – Rauschen mit Ton“ und „tonhaltig – nicht tonhaltig“ vorgenommen.

Die THK-Skala wird verwendet um zu prüfen, wie die Probanden mit dem Begriff der THK umgehen können. Dazu wird vor der Bewertung die THK wie in Abbildung 4.8 definiert.

Die Rangordnung erfolgt anhand der Benutzeroberfläche, die in Abbildung C.2 dargestellt ist. Die Ergebnisse der Rangordnung zeigt Abbildung 3.4.

Für die Skala „Ton – Rauschen“ zeigt sich das erwartete Ergebnis (Diagramm A). Die Stimuli werden entsprechend ihrer THK entlang der Skala angeordnet. Die Stimuli 5 und 6 werden dabei gleich bewertet, da die Probanden nicht in der Lage sind, diese Stimuli zu differenzieren. Das zeigt sich bereits in der Abbildung des Wahrnehmungsraumes, in der die betreffenden Stimuli unmittelbar nebeneinander angeordnet sind.

Ein vergleichbarer Zusammenhang ergibt sich bei der Bewertung der THK, Diagramm C. Die Stimuli 5 und 6 sind diesmal in der erwarteten Reihenfolge. Allerdings kann aufgrund der geringen Probandenzahl und wegen Verwendung des Medians daraus keine Aussage abgeleitet werden.

Zwei Probanden kennzeichnen den reinen Ton (Stimulus 1) als am wenigsten tonhaltig (Rest aufsteigend), einer davon zusätzlich den Stimulus 2 als am zweitwenigsten. Auf Nachfrage geben die Probanden an, dass ein reiner Ton nicht entlang der Skala angeordnet werden kann, da der Ton „in nichts enthalten ist“²³. Daraus folgt, dass ein reiner Ton nicht zwangsläufig als maximal tonhaltig wahrgenommen wird, wie es aus der Norm folgen würde.

In Abbildung 3.4, Diagramm C, sind diese Urteile daher nicht abgetragen, sondern durch ein Ausrufezeichen markiert.

²³ Das sind die gleichen Probanden, die in Diagramm B die Stimuli 1 und 2 „vertauschen“.

3. Perzeptuelle Aspekte der Tonhaltigkeit

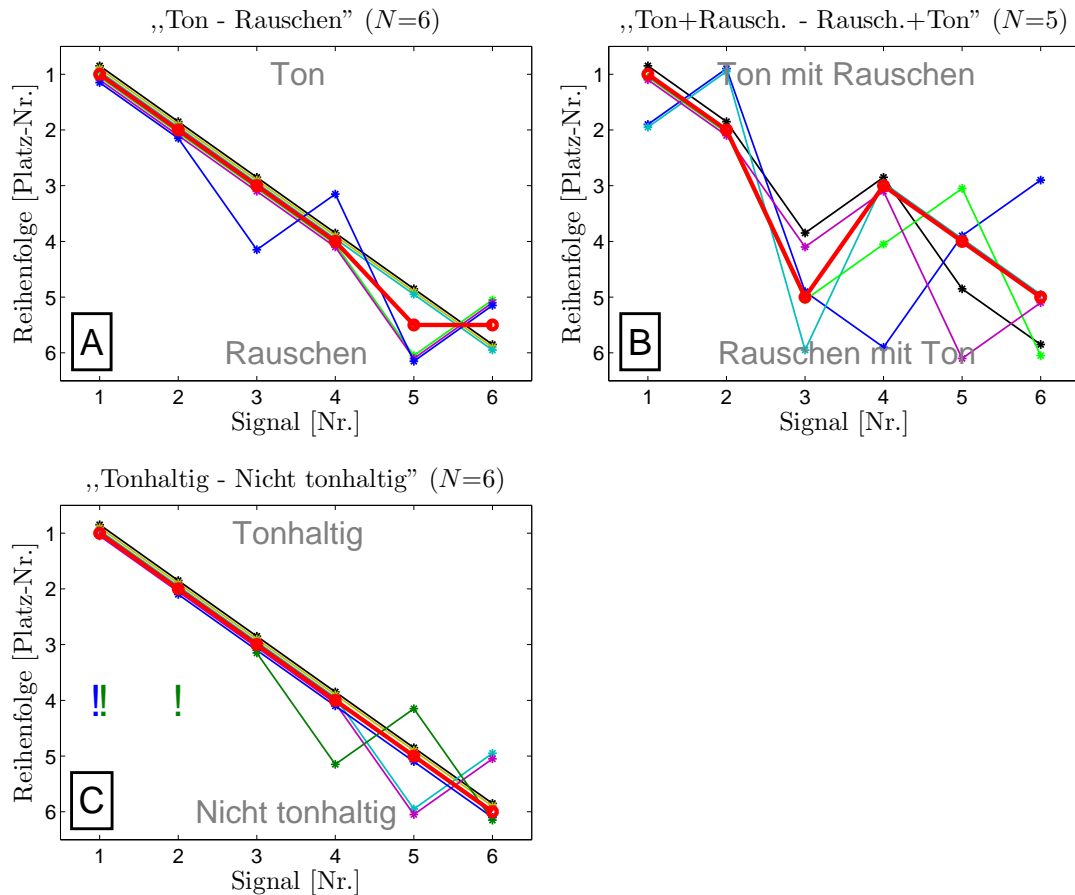


Abbildung 3.4.: Individuelle Rangreihen der Stimuli (Elemente) anhand ausgewählter Kontrastpaare (Konstrukte). Der Median für die Teilnehmer ist als dicke rote Linie dargestellt. Zur besseren Übersicht sind die Daten jeweils leicht entlang der Ordinate verschoben.

Die Skala, anhand derer die Rangordnung in Diagramm B vorgenommen wird, lässt sich nicht direkt aus den Konstrukten, vgl. Tabelle 3.2, ableiten. Die von den Probanden konstruierten Kontrastpaare enthalten nie die verwendete Gegenüberstellung, sondern immer nur Vergleiche der drei Eigenschaften „Ton“, „Rauschen“ oder „Ton+Rauschen“.

Um zu prüfen, ob „Ton mit Rauschen – Rauschen mit Ton“ eine sinnvolle Skala ist, wird sie hier trotzdem verwendet. Aus den Daten zeigt sich allerdings, dass dies nicht der Fall ist. Kein Teilnehmer ist in der Lage, die THK anhand der Skala in eine nachvollziehbare Reihenfolge anzuordnen.

Offenbar wird der Wahrnehmungsraum in erster Linie anhand kategorialer Entscheidungen konstituiert. „Ton“, „Rauschen“ oder „Ton+Rauschen“ sind

entweder vorhanden, oder eben nicht. Die Abstufung innerhalb der Kategorien wird nur dann vorgenommen, wenn mehrere Stimuli der gleichen Kategorie unterschieden werden sollen.

3.3.5. Skalierung der Elemente (Stimuli)

Anhand der im Anhang dargestellten Benutzeroberfläche, Abbildung C.3, wird eine Skalierung ausgewählter Kontrastpaare durchgeführt. Die Ergebnisse dazu sind in Abbildung 3.5 dargestellt.

In der Abbildung sind einerseits Mittelwerte (d.h. die zentrale Tendenz auf einer Intervallskala) und andererseits Mediane (d.h. die zentrale Tendenz auf einer Ordinalskala) inklusive der zugehörigen Streuungsmaße über die Teilnehmer eingetragen.

Diese Vorgehensweise ist ungewöhnlich und wird hier deshalb angewendet, da keiner der Teilnehmer Erfahrungen mit Skalierungen besitzt. Aus diesem Grund kann nicht zweifelsfrei sicher gestellt werden, dass alle Teilnehmer wie erhofft die Skalen in der Benutzeroberfläche (Abbildung C.3) auf Intervallskalenniveau benutzen.

Unter der Annahme von Intervallskalenniveau können die Mittelwerte verwendet werden, bei konservativer Betrachtung auf Ordinalskalenniveau nur die Mediane. Wie aus Abbildung C.3 hervorgeht, sind die Ergebnisse für beide Betrachtungsweisen in allen Fällen vergleichbar. Es ist somit angemessen, wie voraus gesetzt, Intervallskalenniveau für die Skalierungsdaten anzunehmen.

Es werden die gleichen Kontrastpaare wie bei der Herstellung der Rangreihe skaliert. Zusätzlich wird das Kontrastpaar „unangenehm – angenehm“ bewertet, welches die Teilnehmer bei der Evokation der Konstrukte nicht verwenden. Auf diese Weise soll der tatsächliche Einfluss dieses Kontrastpaares ermittelt werden, um zu überprüfen, ob die subjektiven Faktoren wie z.B. die Angenehmheit für die verwendeten Stimuli tatsächlich eine untergeordnete Rolle spielen.

Aus Abbildung 3.5, Diagramm C, ist zu entnehmen, dass die sechs Stimuli sich bzgl. der Angenehmheit kaum voneinander unterscheiden. Es ist somit nachzuvollziehen, dass diese Dimension im Wahrnehmungsraum der Teilnehmer, Abbildung 3.2, keine Rolle spielt. Gleichwohl werden alle Stimuli eher unangenehm als angenehm beurteilt.

Die Ergebnisse für die Skalierung des Kontrastpaares „Ton – Rauschen“ (Abbildung 3.5, Diagramm A) erlauben die gleichen Schlussfolgerungen wie die mittels Rangordnung gewonnenen Ergebnisse, vgl. Abbildung 3.4A. Die Skalierungsdaten passen ebenfalls sehr gut zum ermittelten Wahrnehmungsraum.

Die gleiche Aussage gilt analog für die Bewertung der THK (Abbildung 3.5, Diagramm D). Bemerkenswert ist hier, dass die Teilnehmer, die bei der Herstellung der Rangreihe den reinen Ton als „nicht skalierbar“ bzgl. der THK

3. Perzeptuelle Aspekte der Tonhaltigkeit

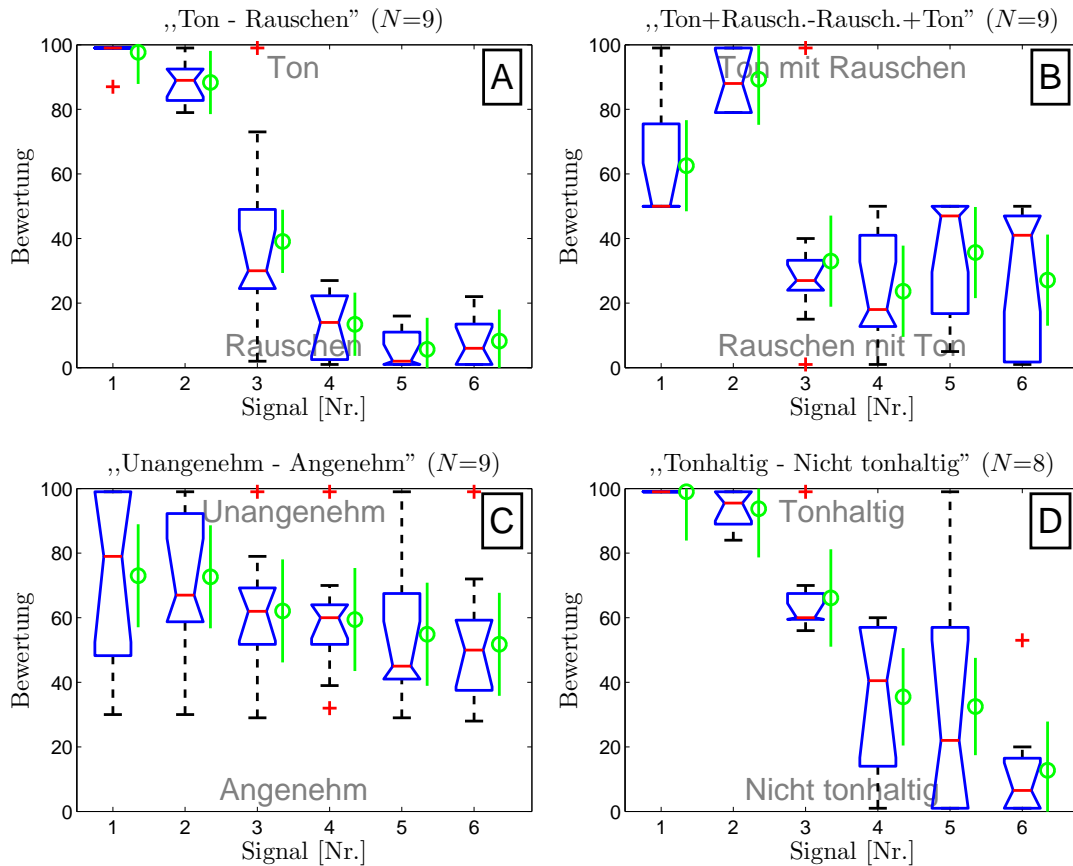


Abbildung 3.5.: Skalierung der Stimuli (Elemente) anhand ausgewählter Kontrastpaare (Konstrukte) auf einer Intervallskala. Dargestellt sind Boxplots nach John W. Tukey [Tuk77], d.h. Median, die beiden Quartilen und die Extremwerte. Die Länge der „Whisker“ beträgt maximal das 1,5-fache des Interquartilabstands und wird immer durch einen Wert aus den Daten bestimmt. „Ausreißer“, d.h. Werte die außerhalb des Whiskers liegen, werden einzeln in das Diagramm eingetragen. Die Einkerbungen („Notches“) geben das Konfidenzintervall für den Median an. Zusätzlich sind in grün die Mittelwerte mit Konfidenzintervall (Bonferroni) angegeben.

klassifiziert haben, den Ton nun doch einheitlich als maximal tonhaltig einstufen.

Ein möglicher Grund dafür könnten die zwischen den Messungen an den Versuchsleiter gerichteten Fragen bzgl. der THK des reinen Tones sein. Es sollte zwar vermieden werden die Teilnehmer zu beeinflussen, in diesem konkreten Fall scheint das aber nicht gelungen zu sein.

Die Skalierung bzgl. des Kontrastpaares „Ton mit Rauschen – Rauschen mit Ton“ ist erneut individuell sehr verschieden, vgl. Abbildung 3.5, Diagramm B. Die Daten sind vergleichbar zu den mit der Rangordnung erhaltenen. Es lässt sich festhalten, dass dieses Kontrastpaar nicht geeignet ist, die Stimuli konsistent daran entlang zu skalieren.

Das ist insofern nicht verwunderlich, als dass dieses Kontrastpaar nicht durch die Probanden selber generiert, sondern zusätzlich miterhoben wird. Im Wahrnehmungsraum der Teilnehmer bilden die Begriffe „Ton mit Rauschen“ und „Rauschen mit Ton“ keine Gegensätze ab.

Als Minimalaussage kann daher nur gefolgert werden, dass die Stimuli 3 und 4 eher als rauschhaft interpretiert werden, während Stimulus 2 als tonal betrachtet wird.

Stimuli 1, 5 und 6 können die meisten Probanden weder der einen noch der anderen Kategorie zuordnen. Bei der Skalierung ordnen sie diesen Stimuli daher einen mittleren Wert zu.

3.4. Diskussion

Die durchgeführten Untersuchungen mit dem Repertory-Grid bringen konsistente Daten hervor und ermöglichen so neue Erkenntnisse zur THK.

Es zeigt sich, dass die Wahrnehmung THK allen Probanden bekannt ist. Da die Probanden so ausgewählt werden, dass sie zuvor keinen Kontakt mit Skalierungsversuchen jeglicher Art hatten und über keine besonderen Kenntnisse der Akustik, Psychoakustik oder THK verfügen, sollte diese Aussage auf beliebige, andere Probanden zu verallgemeinern sein.

Die Probanden verwenden stark individuelle (und daher sehr unterschiedliche) verbale Beschreibungen (Konstrukte) für die THK. Ihr Wahrnehmungsraum bzgl. der THK stimmt aber dennoch weitestgehend überein. Daher können die Konstrukte durch entsprechende Kodierung vereinheitlicht und zusammen gefasst werden. Auf diese Weise ergibt sich eine konsistente, verbale Beschreibung der THK, wie sie auch durch Experten bzgl. der THK vorgenommen würde.

Der umgekehrte Fall ist deutlich aufwändiger. Wenn Probanden ohne jegliche Erfahrungen zur THK die entsprechende Wahrnehmung verdeutlicht werden soll, so steht eine adäquate Beschreibung meist nicht zur Verfügung. Es ist daher problematisch eine verbale Definition zu verwenden (Dilemma der Ambiguität).

Um trotzdem das Perzept THK zu verdeutlichen, können recht einfach akustische Beispiele dargeboten werden. Dieser Vorgang wird im Folgenden als „Orientierung“ bezeichnet, geeignete Beispiele für die Orientierung werden in Kapitel 4 untersucht. Auf diese Weise sind die Probanden in der Lage, sich anhand der Orientierungsgeräusche ihre interne Repräsentation der THK bewusst zu werden und abzurufen.

3. Perzeptuelle Aspekte der Tonhaltigkeit

Anhand der zuvor gewonnenen Erkenntnisse wird die THK als *psychoakustische Empfindungsgröße* nach Zwicker [Zwi82] angesehen. Psychoakustische Empfindungsgrößen werden derart definiert, dass sie getrennt von anderen (psychoakustischen) Größen wahrgenommen und beurteilt werden können. Dies trifft auf die THK auch für Personen zu, die zuvor keine Kenntnisse der THK hatten.

Trotzdem scheint die THK keine „natürliche“ Empfindung, wie zum Beispiel die Tonhöhe oder die Lautheit, zu sein. Für die letztgenannten Empfindungen existieren im Alltagsgebrauch verbale Beschreibungen (laut/leise bzw. hoch/tief) die allgemein gültig und wohldefiniert sind. Das ist nicht der Fall für die THK.

Diese Tatsache kann auf zwei Arten interpretiert werden. Entweder die Bedeutung der THK ist im Alltag nicht groß genug, sodass keine verbale Beschreibung dafür erforderlich ist. Oder die THK kann bereits durch Ausdrücke, die i.A. nicht primär der THK zugeordnet werden, ausreichend beschrieben werden. Dies trifft vor allem auf die in der Gruppe 3 „Ausprägung“ der Tabelle 3.2 genannten Kontrastpaare zu.

Die dort verwendeten Konstrukte sind einerseits der Lautheitswahrnehmung entliehen (laut/leise, starker/schwacher Ton) und beziehen sich andererseits auf eine Trennung in „Figur und Hintergrund“. Das zweite Phänomen entstammt der Gestaltpsychologie und wird auch von Hansen [Han10] berichtet.

Die THK existiert nur dann, wenn sowohl eine Figur (der tonale Anteil) als auch ein Hintergrund (das Rauschen) mehr oder weniger gleichzeitig vorhanden sind (Stimuli 2, 3, 4 in Abbildung 3.2). In diesem Fall entspricht die Wahrnehmung der THK weitestgehend ihrer technischen Definition anhand der maßgeblichen Differenz ΔL nach der Norm.

Für sehr große perzeptive Unterschiede zwischen den Stimuli zeigt sich eine weitere Dimension im Wahrnehmungsraum. Die Klassifizierung findet dann eher kategorial (Ton = „Figur“, Rauschen = „Hintergrund“, Ton im Rauschen = „Figur im Hintergrund“) als auf einer kontinuierlichen Skala statt. Der Wahrnehmungsraum wird in diesem Fall durch 2 Dimensionen aufgespannt, in dem sich die technische THK als Untermenge wiederfinden lässt.

Durch geeignete Instruktion sind die Probanden in der Lage, ihre interne Wahrnehmung der THK zu extrapolieren und auf diese Weise an die technische Definition der THK laut Norm anzunähern. Das zeigt sich anhand der durchgeführten Bewertungen der Stimuli, vgl. z.B. Abbildung 3.4, Diagramm C und Abbildung 3.5, Diagramm D.

4. Wahl eines Orientierungsschalls für die THK

4.1. Motivation und Ziel der Untersuchung

Das Ziel der vorliegenden Untersuchung besteht darin, geeignete Schalle zu identifizieren, die für eine *Orientierung* bzgl. der THK geeignet sind. Orientierung bedeutet in diesem Zusammenhang, Probanden bzw. Beurteiler anhand eines repräsentativen Beispiels das *Perzept* THK zu demonstrieren. Eine ausführliche Diskussion der Problemstellung wird im Kapitel 3 (perzeptuelle Aspekte der THK) geführt.

Eine Orientierung geht über eine Instruktion, d.h. eine rein verbale Beschreibung, deutlich hinaus und bietet auf diese Weise einen erweiterten Zugang zu den vorhandenen akustischen Erfahrungen der Probanden. So kann der Spielraum für individuelle Interpretationen verringert und das Beurteilungskriterium bzgl. der THK vereinheitlicht werden.

4.2. Probanden

Die Untersuchung wird mit drei Gruppen normalhörender Probanden durchgeführt, vgl. Tabelle 4.1.

Tabelle 4.1.: Übersicht der Probanden für die drei Gruppen

Gruppe	Anzahl	weiblich	männlich	Alter in Jahren
Experten 1	20	3	17	24 – 43, Median 29
Experten 2	16	3	13	24 – 43, Median 29
Laien	16	7	9	21 – 50, Median 24

Die erste Gruppe besteht aus 20 Personen aus dem Umfeld des Graduiertenkolleg Psychoakustik an der Universität Oldenburg. Diese Gruppe hat theoretische und praktische Erfahrungen mit der psychoakustischen Bewertung von Schallen, daher wird sie als „Experten 1“ bezeichnet. 16 Personen dieser Gruppe wiederholen den kompletten Ablauf einen Tag später, sie werden als „Experten 2“ zusammengefasst.

Die eigentliche Zielgruppe dieser Untersuchung besteht aus unerfahrenen Probanden, die keine Kenntnisse zum Thema THK und Akustik besitzen. Sie werden im weiteren als „Laien“ bezeichnet und aus der Studentenschaft der Universität Oldenburg gegen eine geringe Aufwandsentschädigung rekrutiert.

Alle Probanden nehmen freiwillig an der Untersuchung teil.

4.3. Stimuli

Es werden vier Stimuli untersucht, die potentiell als Orientierungsschall für die THK geeignet sind. Die Anforderungen an die Stimuli sind zuvor durch Pilotstudien getestet worden und können wie folgt zusammengefasst werden:

- Das Perzept THK soll über einen maximalen Bereich, von sehr schwach bis sehr stark, variieren.
- Die Schalle sollen so weit wie möglich frei von Assoziationen mit bekannten Geräuschen sein. Daher sind künstlich erzeugte Signale zu bevorzugen.
- Die Schalle sollen möglichst elementar sein.
- Die Anzahl gleichzeitig veränderter Empfindungsgrößen soll minimal sein.

Anhand obiger Kriterien werden die Schalle für diese Untersuchung aus einem Sinuston mit einem braunen Rauschen als Hintergrund synthetisch hergestellt. Braunes Rauschen ist einfach zu erzeugen und in grober Näherung vergleichbar zum Spektrum vieler Umweltschalle²⁴.

Für den Sinuston wird eine Frequenz von 700 Hz ausgewählt²⁵. Die Dauer der vier Stimuli beträgt 30 s. Die Schalle unterscheiden sich vor allem durch die Art und Weise, wie Ton und Rauschen sich über die Zeit verändern, vgl. Abbildung 4.1.

Orientierungsschall 1: Der Schalldruckpegel des Rauschens (blaue Kurve) ist während der gesamten Signaldauer konstant. Er wird in der Abbildung willkürlich zu 0 dB gesetzt.

Die Schalldruckamplitude des Tones steigt während der ersten 15 s linear an und fällt während der zweiten Hälfte des Signals linear auf den Ausgangswert zurück. Dieser Verlauf ist in Abbildung 4.1 als Schalldruckpegel (rote Kurve) dargestellt. Das Maximum des Schalldruckpegels für den Ton liegt in etwa 14 dB über dem konstanten Schalldruckpegel des Rauschens. Die Maskierungsschwelle des Rauschens, bezogen auf den verwendeten Ton, wird durch die grüne Kurve angegeben. Liegt der Ton (rote Kurve) oberhalb der Maskierungsschwelle (grüne Kurve) ist er wahrnehmbar, ansonsten wird er durch das Rauschen (blaue Kurve) verdeckt. Somit ist

²⁴ Für eine weitergehende Betrachtung verschiedener Rauschtypen siehe Anhang A.1.

²⁵ Für eine ausführliche Motivation dieser Frequenz siehe Abschnitt 6.2.2.

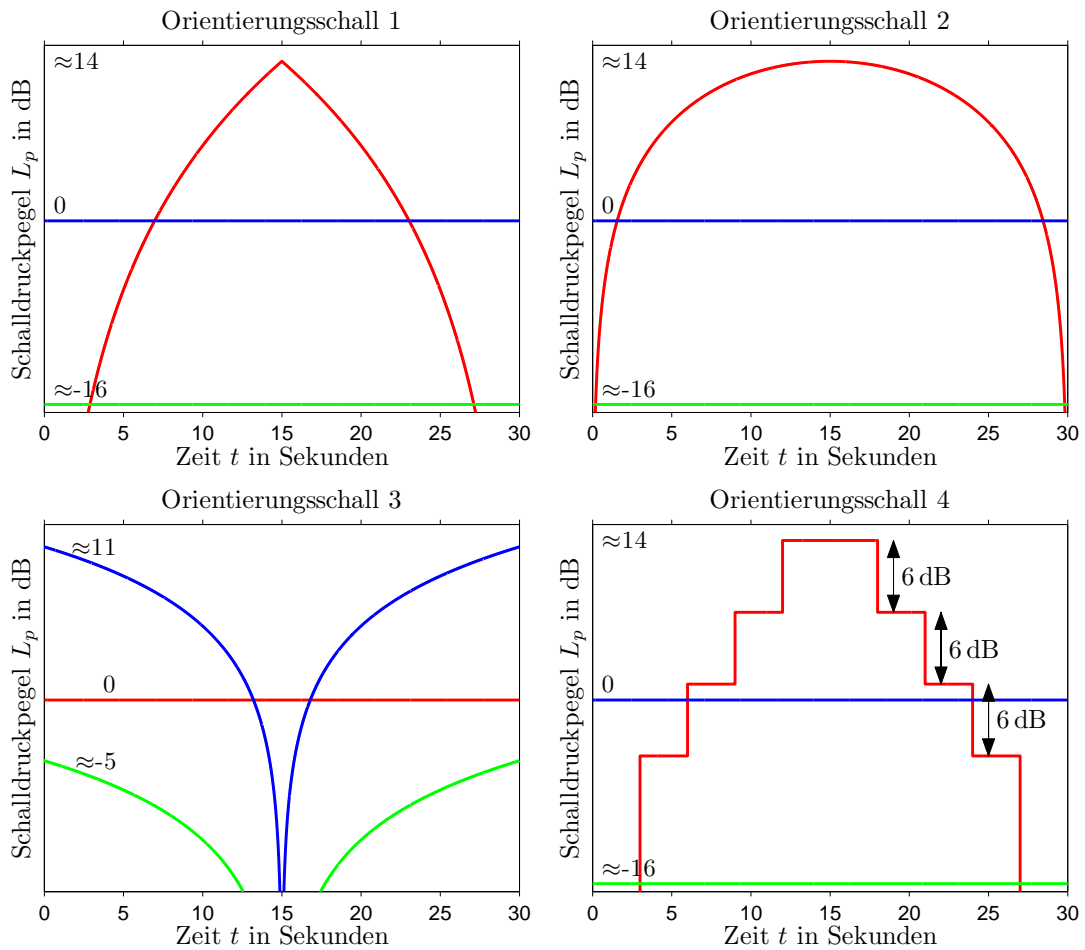


Abbildung 4.1.: Schematische Darstellung der Zeitverläufe des Schalldruckpegels für die potentiellen Orientierungsschalle: rote Kurven symbolisieren den Schalldruckpegel des Sinustones, blaue Kurven den Schalldruckpegel des Hintergrundrauschens und grüne Kurven die Maskierungsschwelle. Liegt der Ton (rote Kurve) oberhalb der Maskierungsschwelle (grüne Kurve) ist er wahrnehmbar, ansonsten wird er durch das Rauschen (blaue Kurve) verdeckt.

aus dem Diagramm zu entnehmen, dass am Anfang und Ende des Signals der Ton etwa 3s nicht hörbar ist.

Orientierungsschall 2: Der Schalldruckpegel des Rauschens (blau) ist erneut konstant. Die konstante Schalldruckamplitude des Tones wird mit einem *Hanning-Fenster*²⁶ mit einer Länge von 30s multipliziert, um ein stetiges Ein-

²⁶ Auch als „Raised-Cosinus Fenster“ bezeichnet, siehe z.B. [BT59].

und Ausschalten des Tones zu erreichen.

Der Ton wird die ersten 15 s lauter, erreicht im Bereich um 15 s einen Schalldruckpegel (rot) von etwa 14 dB über dem Schalldruckpegel des Rauschens und sinkt danach wieder ab.

Verglichen mit Orientierungsschall 1 ändert sich der Schalldruckpegel für dieses Signal zu Anfang und am Ende schneller, in der Mitte jedoch langsamer. Der Ton ist zu Beginn und am Ende des Schalls zwar schwach, aber wahrnehmbar.

Orientierungsschall 3: Der Schalldruckpegel des Tones (rot) ist konstant, die Schalldruckamplitude des Rauschens wird in den ersten 15 s linear abgesenkt und während der zweiten Hälfte des Signals auf die gleiche Weise wieder angehoben. Dieses Verhalten ist im Diagramm als Schalldruckpegel (blau) dargestellt.

In der Mitte des Schalls, bei etwa 15 s, ist das Rauschen kurz gar nicht zu hören, d.h. der Ton ist dort nicht maskiert und maximal wahrnehmbar. Zu Beginn und am Ende des Schalls ist der Ton ebenfalls deutlich wahrnehmbar.

Orientierungsschall 4: Der Schalldruckpegel des Rauschens (blau) ist konstant. Der Schalldruckpegel des Tones (rot) wird zwischen 3 s und 15 s alle drei Sekunden um 6 dB angehoben, d.h. die Schalldruckamplitude wird verdoppelt. Ab 15 s wird der Schalldruckpegel des Tones dazu symmetrisch abgesenkt. Das Maximum des Schalldruckpegels für den Ton liegt, wie für den Orientierungsschall 1 und 2, in etwa bei 14 dB über dem konstanten Schalldruckpegel des Rauschens.

Am Anfang und Ende des Schalls ist der Ton etwa 3 s nicht wahrnehmbar. Der Orientierungsschall 4 ist dem Orientierungsschall 1 sehr ähnlich, der wesentliche Unterschiede besteht darin, dass für den Schall 4 die Pegeländerung in diskreten Schritten erfolgt, während sie bei Schall 1 kontinuierlich stattfindet.

4.4. Durchführung

Die vier Schalle werden mittels lateinischer Quadrate in ausbalancierter Reihenfolge präsentiert. Die Fragestellung an die Probanden lautet: „Für wie geeignet halten sie die dargebotenen Geräusche, den Höreindruck THK Personen zu demonstrieren, welche die Definition dieses Begriffes nicht kennen?“

Die Schalldarbietung erfolgt über elektrostatische Kopfhörer Sennheiser HE60 mit der Versorgungseinheit HEV70 bei gleicher Lautheit von 10 sone in einer schallisolierten Kabine. Die Aufgabe der Probanden besteht darin, die Schalle mittels einer 4-stufigen Kategorienskala (siehe Abbildung 4.2) zu bewerten. Für

die Bewertung haben die Probanden 6 s Zeit, bevor automatisch der nächste Schall präsentiert wird.

Die Experten führen den Versuch zwei mal mit mindestens einem Tag Abstand durch. Ihnen werden vorab keine Schalle zur Orientierung dargeboten, da sie bereits Kenntnisse zum Thema THK besitzen.

Den Laien werden die vier zu bewertenden Schalle zwei mal hintereinander in unterschiedlicher Reihenfolge vorgespielt. Der erste Durchgang dient lediglich der Orientierung, die Beurteilung erfolgt jeweils nur für die zweite Darbietung.

nicht besonders geeignet	einigermaßen geeignet	gut geeignet	sehr gut geeignet
(1)	(2)	(3)	(4)

Abbildung 4.2.: Verwendete Kategorienskala zur Bewertung der Orientierungsschalle. Die Zahlen dienen der Kodierung der Kategorien für die weitere Auswertung; sie werden den Probanden während der Beurteilung nicht dargeboten.

4.5. Bewertungsergebnisse

Im Weiteren werden lediglich gruppenweise Auswertungen der gewonnenen Daten vorgenommen und dargestellt. Der Vollständigkeit halber sind die individuellen Daten im Anhang [D](#) in der Tabelle [D.2](#) zusammengefasst.

Die Probanden aller drei Gruppen beurteilen die dargebotenen Schalle als prinzipiell geeignet, um zur Orientierung für den Höreindruck THK verwendet zu werden. Vergleiche dazu [Abbildung 4.3](#): mehr als 80% der abgegebenen Urteile liegen in den Kategorien 2, 3 und 4, sind also mindestens „einigermaßen geeignet“ oder besser.

Alle drei Gruppen verwenden die vier zur Verfügung stehenden Kategorien in etwa gleich häufig. Es existieren nur geringe, unsystematische Unterschiede zwischen den Gruppen.

Die Auswertung erfolgt im Weiteren auf ordinalem Skalenniveau, da bei der Entwicklung der Bewertungsskala ([Abbildung 4.2](#)) nicht geprüft wurde, ob die vier Kategorien jeweils den gleichen Abstand zueinander haben. Daher werden zunächst gruppenweise die Häufigkeiten der verwendeten Kategorien für jeden Schall aufgetragen, siehe [Abbildung 4.4](#).

Aufgrund des ordinalen Skalenniveaus können im Weiteren nur parameterfreie statistische Verfahren angewendet werden. Daher werden die kategorial

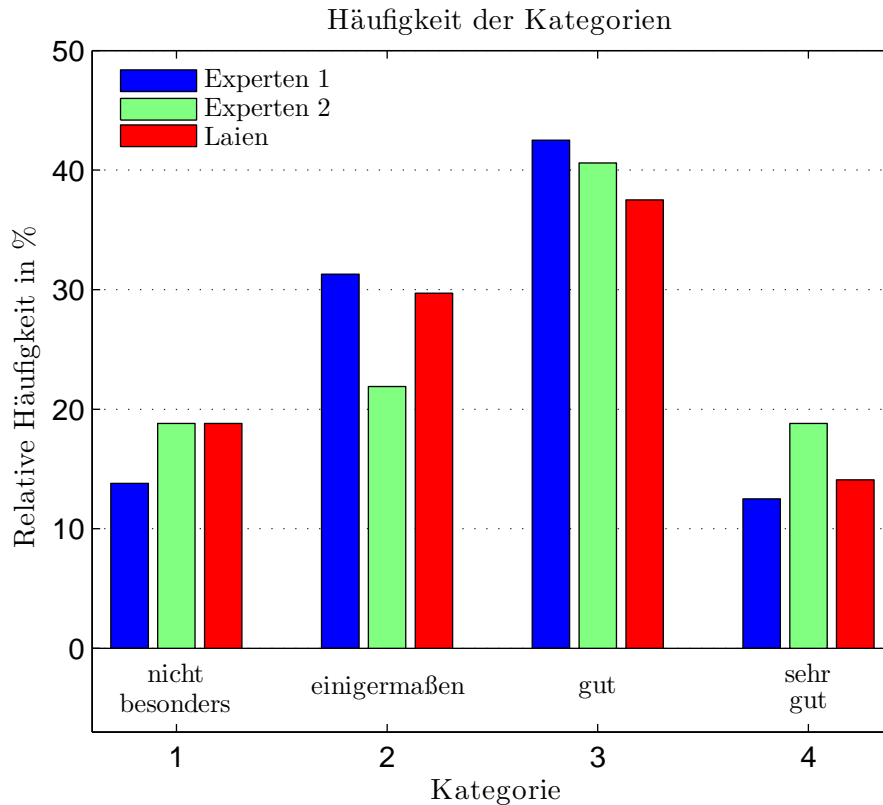


Abbildung 4.3.: Häufigkeit der verwendeten Kategorien, getrennt für Laien und Experten.

erhobenen Urteile zunächst für jeden Probanden individuell in Ränge umgerechnet. Diese werden anhand der Zahlenwerte der Urteile (vgl. Abbildung 4.2) mit aufsteigender Größe vergeben. Bei Rangbindungen werden die Ränge der gebundenen Urteile durch arithmetische Mittelung berechnet. Das Resultat zeigt Abbildung 4.5.

Um eine erweiterte Datenanalyse durchführen zu können, werden die ordinal skalierten Daten mittels des *Law of Categorical Judgement* (vgl. z.B. [BD06]) auf Intervallskalenniveau transformiert. Das Verfahren ist im Anhang D.4 näher erläutert und für die vorliegenden Daten durchgeführt. Das Ergebnis der Transformation zeigt Abbildung 4.6.

Die Unterschiede zwischen den potentiellen Orientierungsschallen können anhand dieser Darstellung genauer quantifiziert und insbesondere direkt verglichen werden.

Im folgenden werden post-hoc die Unterschiede sowohl zwischen den Gruppen als auch zwischen den Orientierungsschallen auf Signifikanz geprüft²⁷.

²⁷ Für alle statistischen Auswertungen wird die Statistik-Software SPSS [SPS02] benutzt.

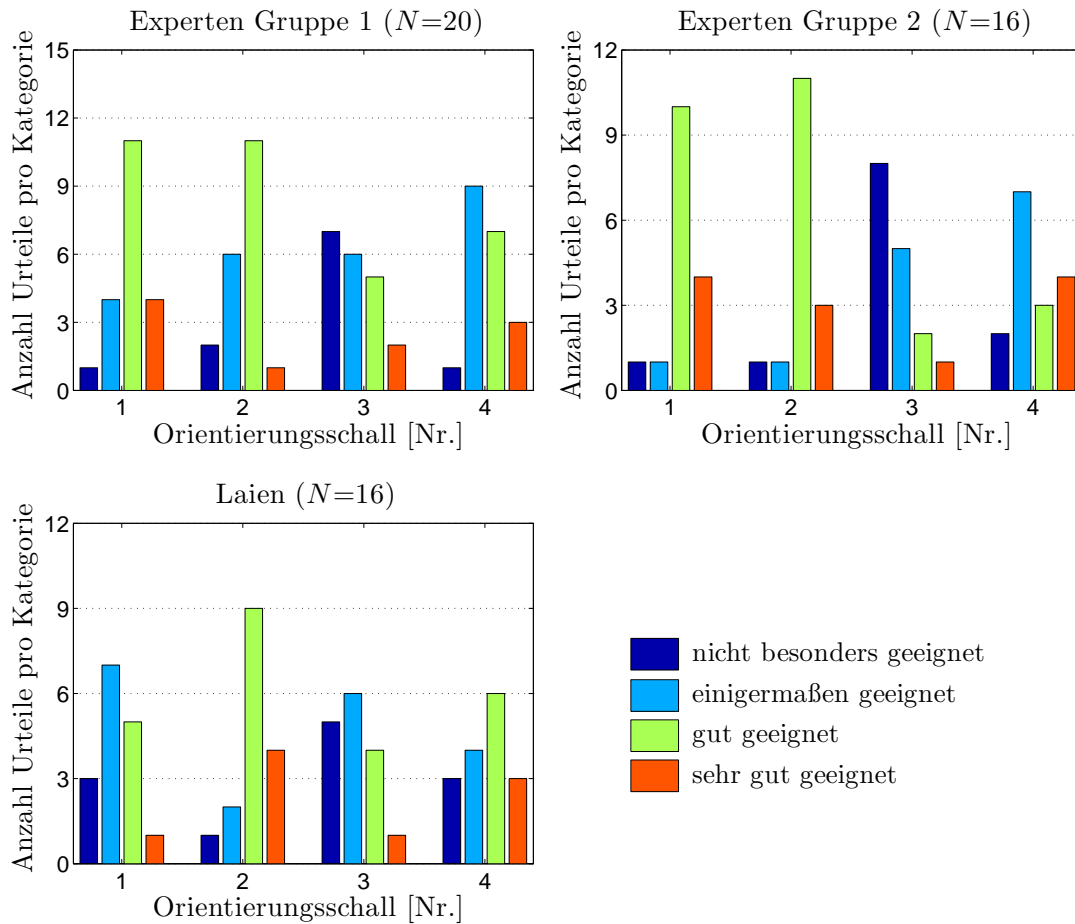


Abbildung 4.4.: Häufigkeit der verwendeten Kategorien, aufgeteilt nach Schallen für jede der drei Gruppen. Das Maximum der Ordinaten in den drei einzelnen Diagrammen entspricht jeweils 75% der Anzahl der Teilnehmer (N).

Zunächst werden die drei Gruppen mittels Kruskal-Wallis-Test verglichen. Für Orientierungsschall 1 zeigt sich ein signifikanter Effekt²⁸ ($\chi^2(2, N=52)=8.63$, $p=.011$) der Gruppen auf die Urteile. Für die drei anderen Orientierungsschalle ist hingegen kein Effekt der Gruppen nachweisbar.

Da der Kruskal-Wallis-Test hier vier mal (für jeden Orientierungsschall einmal) durchgeführt wird, muss die α -Fehler-Inflation berücksichtigt werden. Hier wird die Bonferroni-Methode angewendet, wodurch die Signifikanz für Orientierungsschall 1 auf $p=.044$ absinkt. Trotzdem zeigt sich weiterhin ein Effekt der Gruppe auf einem Signifikanzniveau von 5%.

Dieser statistisch bedeutsame Unterschied wird offenbar dadurch bedingt,

²⁸ Die Berechnung erfolgt exakt, d.h. nicht durch asymptotische Näherung.

4. Wahl eines Orientierungsschalls für die THK

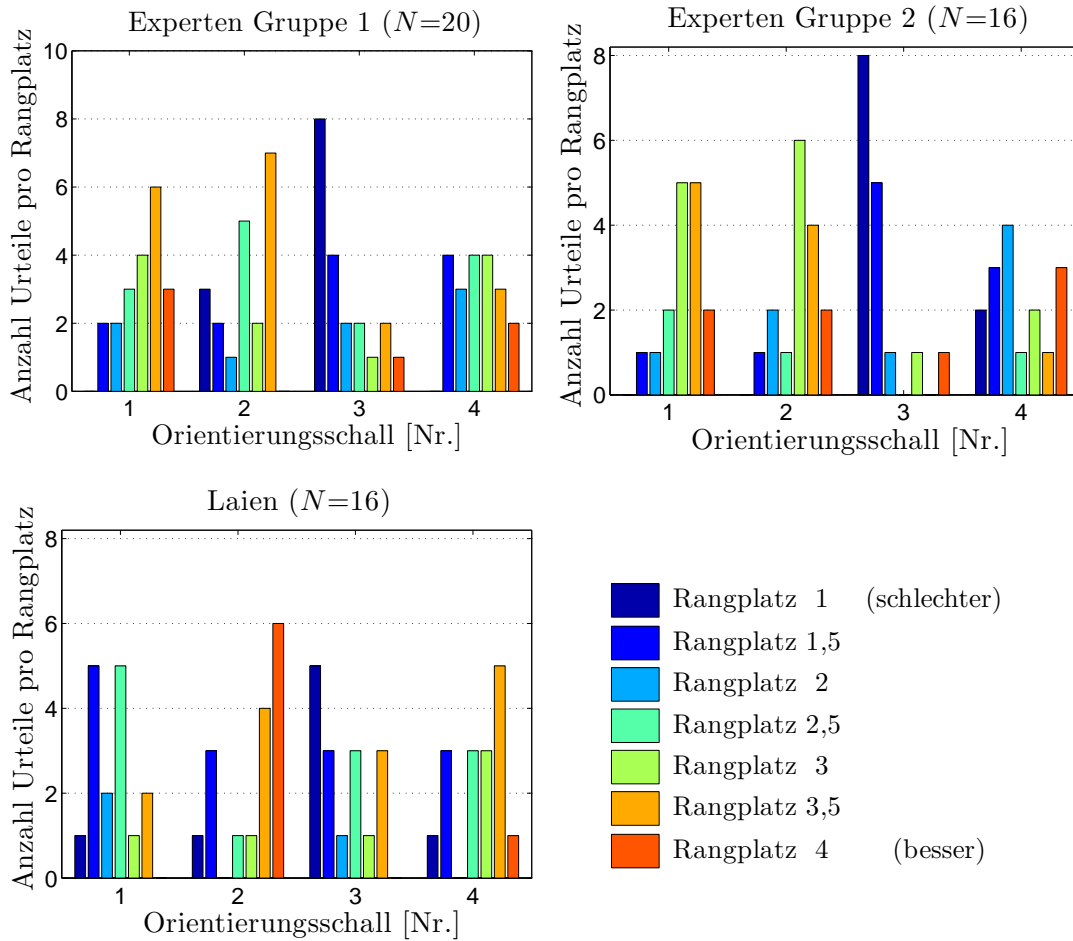


Abbildung 4.5.: Häufigkeit der Ränge für die potentiellen Orientierungsschalle, nach Gruppen getrennt. Die Ränge werden individuell für jeden Probanden einzeln aus seinen kategorialen Bewertungen ermittelt. Das Maximum der Ordinaten in den drei einzelnen Diagrammen entspricht jeweils 75% der Anzahl der Teilnehmer N .

dass die Laien Orientierungsschall 1 deutlich schlechter beurteilen als die Experten, siehe Abbildung 4.6. Die drei anderen Orientierungsschalle werden ebenfalls leicht unterschiedlich bewertet, der Effekt ist jedoch nicht statistisch signifikant.

Um diese Hypothese weiter zu untersuchen, wird die Übereinstimmung der beiden Experten-Gruppen untersucht. Dazu wird der Mann-Whitney-U-Test getrennt für die vier Orientierungsschalle angewendet.

Streng genommen müsste dieser Test für verbundene Stichproben durchgeführt werden, da die gleichen Probanden in den beiden Experten-Gruppen 1 und 2 teilnehmen. Da die Urteilsabgabe jedoch anonym erfolgt, können die

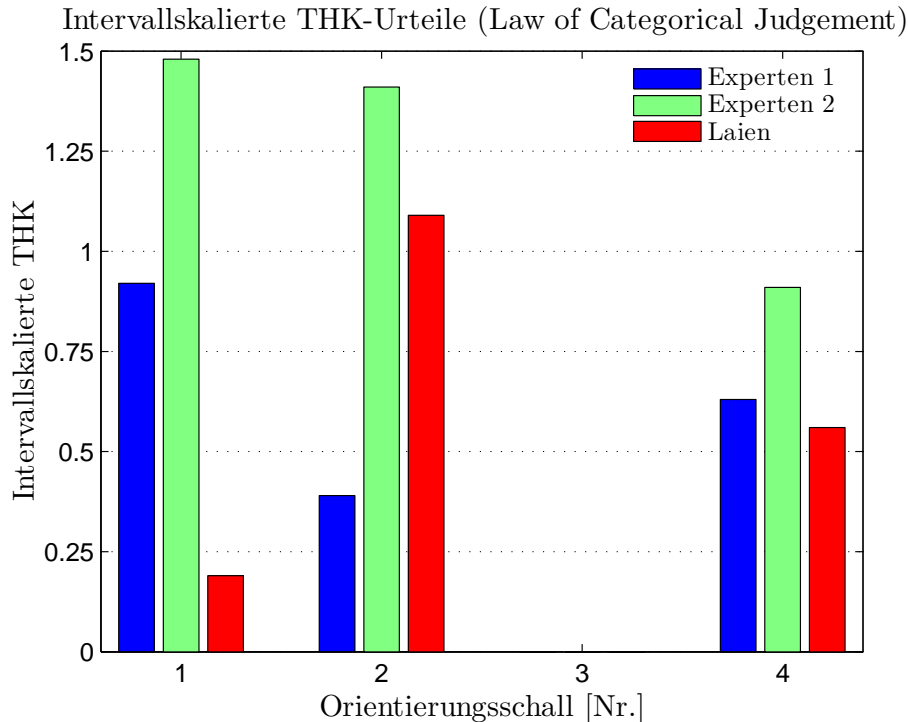


Abbildung 4.6.: Anhand des *Law of Categorical Judgement* auf Intervallskalenniveau transformierte Probandenurteile für die drei Probandengruppen.

identischen Probanden zwischen den beiden Gruppen nicht korrekt „verbunden“, d.h. zugeordnet werden. Es bleibt somit nur die Möglichkeit, diese beiden Gruppen als unabhängig einzustufen und den Test für unabhängige Stichproben durchzuführen.

Für Orientierungsschall 2 ergibt sich dann eine unkorrigierte, exakte, zweiseitige Signifikanz von $p=.058$. Für die drei anderen Orientierungsschalle ist die Signifikanz deutlich geringer. Bei Berücksichtigung der α -Fehlerkorrektur (Bonferroni) wird ersichtlich, dass es keine signifikanten Unterschiede zwischen den beiden Gruppen Experten 1 und Experten 2 gibt. Für Orientierungsschall 2 existiert im unkorrigierten Fall allenfalls ein tendenzieller Unterschied.

Es ist somit gerechtfertigt, die beiden Experten-Gruppen gemeinsam weiter auszuwerten. Das wird realisiert, indem die Daten der Experten-Gruppe 2 stellvertretend für beide Experten-Gruppen verwendet werden.

Das ist vor allem dadurch motiviert, dass die Experten während ihrer ersten Beurteilung die Stimuli noch nicht kennen und die Bewertungen ausschließlich basierend auf ihrer „internen Referenz“ abgeben müssen.

Im zweiten Beurteilungsdurchgang (Experten, Gruppe 2) sind die Stimuli bereits bekannt und das Urteil wird anhand der Eindrücke bzgl. aller Stimuli neu

verankert. Insofern sind die Beurteilungen der Experten-Gruppe 1 nicht vorbehaltlos mit den Urteilen der Experten-Gruppe 2 vergleichbar. Es ist daher günstiger, die Daten der zweiten Experten-Gruppe stellvertretend für die Experten zu verwenden und die beiden Gruppen nicht zusammen zu fassen.

Zusätzlich entfällt dadurch die Notwendigkeit, einen geeigneten statistischen Test durchzuführen, der berücksichtigt, dass die Probanden aus den beiden Experten-Gruppen zwar identisch sind, aber innerhalb der beiden Gruppen im Nachhinein nicht mehr korrekt zugeordnet werden können.

Um Unterschiede zwischen den Orientierungsschallen zu finden, wird der Friedman-Test für jede Gruppe getrennt durchgeführt.

Für die Laien ergibt sich die exakte Signifikanz zu $p=.039$ ($\chi^2(3, N=16)=8.2$). Für die Experten (Gruppe 2)²⁹ gilt $\chi^2(3, N=16)=17.6$, $p<.001$. Es existieren daher für beide Gruppen gleichermaßen signifikante Unterschiede zwischen den Orientierungsschallen.

Zur Ermittlung, welche Orientierungsschalle sich jeweils statistisch belastbar voneinander unterscheiden, wird der Wilcoxon-Test gruppenweise für alle sechs möglichen Kombinationen der Orientierungsschalle angewendet.

Für die Laien ergeben sich ohne Korrektur der Inflation des α -Fehlers zwei signifikante³⁰ Unterschiede für Orientierungsschalle 1 und 2 ($p=.020$) sowie für Orientierungsschalle 2 und 3 ($p=.027$). Bei Korrektur des α -Fehlers (Bonferroni) ist kein signifikanter Unterschied erkennbar. Die Laien bewerten somit Orientierungsschall 2 allenfalls tendenziell besser als Orientierungsschall 1 und 3.

Die Experten (Gruppe 2) bewerten Orientierungsschall 3 ohne Korrektur des α -Fehlers auf dem Signifikanzniveau von 5% signifikant³⁰ schlechter als alle anderen Schalle. Bei Bonferroni-Korrektur bleibt der Unterschied signifikant³⁰ zwischen Orientierungsschall 1 und Orientierungsschall 3 ($p=.009$) sowie zwischen Orientierungsschall 2 und Orientierungsschall 3 ($p=.010$)³¹.

Die Daten zeigen, dass Orientierungsschall 3 für alle Gruppen (signifikant) am schlechtesten ist. Die Bestimmung des besten Orientierungsschalls ist hingegen nicht trivial. Ein möglicher Grund dafür ist die große Anzahl an ranggebundenen Urteilen: die Experten-Gruppe 1 weist 42,5% Urteile mit Rangbindungen auf, die zweite Experten-Gruppe 37,5% und die Laien 62,5%.

Die große Anzahl ranggebundener Urteile resultiert vor allem aus der Tatsache, dass die vier potentiellen Orientierungsschalle auf einer 4-stufigen Katego-

²⁹ Für die Experten Gruppe 1 zeigt der Friedman-Test ($\chi^2(3, N=20)=11.2$) eine Signifikanz von $p=.009$.

³⁰ exakt, zweiseitig

³¹ Vollständigkeitshalber sei erwähnt, dass die Experten-Gruppe 1 signifikante Unterschiede (unkorrigiert, zweiseitig) zwischen Schall 1 und 3 ($p=.006$) sowie zwischen Schall 3 und 4 ($p=.052$) aufweisen. Bei Bonferroni-Korrektur bleibt ein signifikanter Unterschied ($p=.036$) zwischen Schall 1 und 3 erhalten.

rialskala bewertet werden. Anhand der diskutierten Skala ist es den Probanden jedoch nicht möglich, die Unterschiede der vier verwendeten Orientierungsschalle adäquat abzubilden.

Unter der Voraussetzung gleicher Abstände zwischen den Kategorien der verwendeten Skala (Abbildung 4.2) kann für die Probanden-Urteile Intervallskalenniveau angenommen werden. Es ist somit möglich, die statistische Analyse auch anhand von parametrischen Testverfahren durchzuführen.

Die entsprechenden Ergebnisse sind im Anhang, Kapitel D.5, aufgeführt. Sie bestätigen die zuvor dargestellten Befunde, offenbaren aber selbst keine neuen Erkenntnisse. Daher wird an dieser Stelle nicht weiter darauf eingegangen.

4.6. Diskussion

Das vorrangige Ziel dieser Untersuchung besteht darin, den besten der vier zur Auswahl stehenden Orientierungsschalle zu bestimmen. Der beste Schall wird nachfolgend als der von den Probanden präferierte Schall angesehen.

Offensichtlich haben einige Probanden keine eindeutige Präferenz bzgl. des am besten geeigneten Orientierungsschalls. Das zeigt sich bei Betrachtung der individuellen Daten, vgl. Tabelle D.2 im Anhang D.3. Die in der Tabelle grau hinterlegten Probanden bewerten drei oder sogar alle vier Schalle als gleich geeignet. Diese Teilnehmer liefern nur bedingt Informationen darüber, welcher Orientierungsschall der Beste ist.

Werden die Probanden ohne eindeutige Präferenz aus der Analyse ausgeschlossen, so wird das Ergebnis eindeutiger, ohne jedoch statistisch signifikant zu werden. Aus diesem Grund wird die Analyse weiterhin für alle Probanden durchgeführt.

Die Laien bevorzugen tendenziell Orientierungsschall 2, gefolgt von Schall 4 und Schall 1. Der potentielle Orientierungsschall 3 ist am wenigsten geeignet. Für die Experten wird nur die Gruppe 2, d.h. der zweite Bewertungsdurchgang der Experten berücksichtigt. Die Daten zeigen, dass die Orientierungsschalle 1 und 2 gleichermaßen präferiert werden, gefolgt von Schall 4. Auch hier wird der potentielle Orientierungsschall 3 für am wenigsten geeignet befunden, er ist signifikant schlechter als alle anderen Schalle.

Aufgrund der gefundenen Daten wird der potentielle Orientierungsschall 2 als Prototyp der THK ausgewählt, um ihn zur akustischen Demonstration der THK zu verwenden und diesen Schall für die Orientierung einzusetzen. Schall 2 wird von den Laien präferiert und von den Experten ebenfalls zu den beiden geeignetsten Schallen gezählt. Der Modus der Beurteilungen liegt für alle Gruppen in der Kategorie „gut geeignet“.

4. Wahl eines Orientierungsschalls für die THK

Die auf Intervallskalenniveau durchgeführte Analyse (ANOVA, Anhang D.5) bestätigt diese Ergebnisse. Der über die Experten (Gruppe 2) und Laien zusammengefasst präferierte, d.h. beste Orientierungsschall ist demnach Orientierungsschall 2 ($M=3,00$ $SD=0,76$), gefolgt von Orientierungsschall 1 ($M=2,66$ $SD=0,90$) und Orientierungsschall 4 ($M=2,56$ $SD=1,01$). Orientierungsschall 3 ($M=1,91$ $SD=0,93$) ist der deutlich am wenigsten präferierte Stimulus.

Möglicherweise präferieren die Laien den Orientierungsschall 2, da der Ton in diesem Schall durchgehend hörbar ist und nicht, wie in Schall 1 und 3, am Anfang und Ende komplett maskiert wird. Für die Experten scheint es keine Rolle zu spielen, ob der Ton temporär maskiert wird oder nicht.

Im Anschluss an die kategorialen Bewertungen der Orientierungsschalle wird ein offenes Interview mit den Laien durchgeführt. Diese Befragung erfüllt nicht die Voraussetzungen, systematisch ausgewertet zu werden. Sie liefert jedoch Hinweise auf die Notwendigkeit, eine Orientierung durchzuführen, bevor subjektive Bewertungen zur THK vorgenommen werden.

Die folgenden Punkte umfassen die wichtigsten Aussagen der Laien (bzw. deren Interpretation), die übereinstimmend in dieser Art geäußert wurden:

- THK ist kein Begriff des täglichen Lebens – sämtliche Laien kennen diesen Begriff vor Teilnahme an dieser Studie nicht.
- Der überwiegende Teil der Laien kennt hingegen sehr wohl die Wahrnehmung bzw. den Höreindruck THK aus dem persönlichen Umfeld.
- Beschreibungen bzw. Erklärungen zur THK sind oft nicht eindeutig. Sie werden je nach Erfahrung der Probanden unterschiedlich aufgefasst und interpretiert. Erst durch eine akustische Orientierung wird das Konstrukt THK einheitlich(er) verstanden.
- Ohne akustische Orientierung betrachten einige Laien ihre Aufgabe zunächst als Detektionsexperiment. Durch die Orientierung wird die Aufgabe leichter verständlich.
- Unabhängig zu den kategorialen Bewertungen äußern die Probanden, dass alle potentiellen Orientierungsschalle geeignet sind, die THK zu demonstrieren. Im Anschluss an die Bewertungen wird Schall 3 hingegen verbal deutlich schlechter eingeschätzt als die anderen Schalle, da er ökologisch weniger valide erscheint.

4.7. Standard-Orientierung für die Tonhaltigkeit

Die aufgeführten Aspekte zeigen, dass Probanden vor subjektiven Bewertungen der THK zunächst sinnvoll instruiert werden müssen. Dazu gehört eine, an die Wahrnehmung des Menschen angepasste, akustische Orientierung. Erst

nachdem das Perzept THK unmissverständlich bekannt ist, sollte das spezifische Ziel der Untersuchung, bzw. die konkrete Aufgabe der Probanden, bekannt gegeben werden.

Es wird daher empfohlen, bei allen THK-Bewertungen mit „Laien“, d.h. Probanden die im Vorfeld nicht wissen, was der Begriff bzw. die Definition „Tonhaltigkeit“ bedeutet, die Standard-Orientierung, wie in Abbildung 4.7 beschrieben, durchzuführen. Auf diese Weise kann ein einheitliches Vorgehen und somit konsistente Messungen der subjektiven THK gewährleistet werden.

1. Kurze schriftliche Definition / Erläuterung der THK
2. Orientierung mit Orientierungsschall 2
3. Bei Bedarf zusätzliche Orientierung mit:
 - a) Umweltgeräuschen
 - b) Schallen aus der folgenden Studie
4. Rücksprache und Möglichkeit zur Nachfrage
5. Durchführung der „Hauptstudie“

Abbildung 4.7.: Standard-Orientierung: Grundlegende Schritte bei der Durchführung von Instruktion und Orientierung für die THK.

Zu 1: Die Orientierung sollte dem Probanden schriftlich vorgelegt werden. Wenn methodisch für den weiteren Ablauf der Studie keine Bedenken bestehen, kann die Orientierung mündlich vom Versuchsleiter ergänzt werden. Der Versuchsleiter sollte dabei jedoch auf keinen Fall Assoziationen bzgl. der THK bei den Probanden hervorrufen, da diese der Intention einer definierten und für alle Probanden einheitlichen Orientierung entgegenstehen. In jedem Fall sollte der Proband die Möglichkeit haben, jederzeit Nachfragen an den Versuchsleiter zu richten.

Eine geeignete Definition der Tonhaltigkeit zeigt Abbildung 4.8.

Definition der Tonhaltigkeit:

Ein Geräusch ist tonhaltig, wenn einzelne oder mehrere Töne erheblich aus dem Geräusch hervortreten und deutlich hörbar sind.

Das gilt auch wenn Klänge, Heulen, Quietschen, Kreischen oder andere tonartige Phänomene aus dem Geräusch hervortreten.

Abbildung 4.8.: Standard-Orientierung: Definition der THK.

Zu 2: Dieser Teil dient vor allem dazu, den Probanden die Empfindungsgröße THK zu demonstrieren. In den meisten Fällen kennt der Proband die

4. Wahl eines Orientierungsschalls für die THK

Wahrnehmung THK, kann sie jedoch nicht dem Begriff THK, bzw. der komplexen Definition der THK zuordnen.

Das Ziel besteht daher darin, die bekannten akustischen Eindrücke des täglichen Lebens mit dem Begriff THK zu verknüpfen.

Der Text zur Vorlage für den Probanden wird aufgrund der Erfahrungen dieser Arbeit wie in Abbildung 4.9 empfohlen.

Um Ihnen einen akustischen Eindruck von der Hörempfindung Tonhaltigkeit zu vermitteln, wird im Folgenden ein Beispiel dargeboten.

Das Beispiel dauert etwa 30 Sekunden. Achten Sie während dieser Zeit auf die Eigenschaft des Signals, die sich am stärksten verändert: dies ist die Tonhaltigkeit.

Die Gesamtlautstärke des Signals verändert sich während der Darbietung ebenfalls leicht. Das ist technisch bedingt und soll nicht weiter beachtet werden.

Abbildung 4.9.: Standard-Orientierung: Darbietung des Orientierungsschalls.

Zu 3: Bei Bedarf und je nach Untersuchungsziel können ergänzende Beispiele präsentiert werden, die weitere Aspekte der THK abdecken. Im Rahmen dieser Arbeit sind gute Erfahrungen mit den Umweltgeräuschen von Pompetzki [Pom98], siehe Anhang B.1, erzielt worden.

Alternative Schalle sind ebenfalls denkbar, z.B. wenn in der nachfolgenden Untersuchung ausschließlich synthetische Signale verwendet werden. Je nach Untersuchungsziel kann es in solchen Fällen hilfreich sein, diese Schalle bereits zur Orientierung zu verwenden.

Die Anweisung für den Probanden kann sinngemäß wie in Abbildung 4.10 erfolgen.

Zu 4: Abschließend soll der Proband über den Erfolg der Orientierung befragt werden und Gelegenheit zur Nachfrage bekommen, siehe Abbildung 4.11.

Zu 5: Nachdem die Orientierung bzgl. der THK erfolgreich durchgeführt worden ist, kann anschließend die Instruktion für die nachfolgenden, spezifischen Messungen zur (subjektiven) THK erfolgen. Dafür wird in den meisten Fällen erneut eine Instruktion der Probanden erforderlich. Eventuell müssen die Probanden darüber hinaus bzgl. ihrer folgenden Aufgaben und Tätigkeiten zusätzlich orientiert werden.

Die Anforderung, ob und wie zu dieser Phase der Studie instruiert bzw. orientiert werden soll, hängt ausschließlich von den nachfolgenden Untersuchungen ab. Die Entscheidung diesbezüglich ist komplett unabhängig von der oben vorgestellten Orientierung zu THK.

Sie hören nun Aufnahmen von typischen, tonhaltigen Umweltgeräuschen.

Achten Sie bitte besonders auf die Tonhaltigkeit dieser Geräusche, das heißt die Ausprägung der tonalen Anteile im Vergleich zum Hintergrundschall. Die Tonhaltigkeit ist umso größer, je „lauter“ und somit ausgeprägter die tonalen Anteile sind.

Die tonalen Anteile können zusätzlich verschiedene Tonhöhen haben, das heißt sie klingen „höher“ oder „tiefer“. Diese Eigenschaft der tonalen Anteile ist für die Bewertung der Tonhaltigkeit unbedeutend. Bitte versuchen Sie daher, die Tonhöhe bei der Bewertung der Tonhaltigkeit nicht zu berücksichtigen.

Abbildung 4.10.: Standard-Orientierung: Darbietung ergänzender Orientierungsschalle.

Sie sollten nun einen akustischen Eindruck von der Hörempfindung Tonhaltigkeit bekommen haben.

Wenn Sie jetzt wissen, was Tonhaltigkeit bedeutet, können Sie die folgenden Versuche leicht durchführen.

Wenn Sie noch unsicher sind, dann informieren Sie bitte den Versuchsleiter darüber.

Abbildung 4.11.: Standard-Orientierung: Abschließende Nachfrage.

Die hier beschriebene Vorgehensweise zur Orientierung bzgl. der THK kommt bei allen im Folgenden präsentierten Untersuchungen zur Anwendung. Dabei muss die Orientierung für jeden Probanden mindestens einmal durchgeführt werden. Es ist jedoch nicht zwingend erforderlich, die Orientierung vor jeder weiteren Untersuchung des selben Probanden zu wiederholen.

4. Wahl eines Orientierungsschalls für die THK

5. Tonhaltigkeit von Einzeltönen in Rauschen im vollständigen Paarvergleich

5.1. Motivation und Ziel der Untersuchung

Zwei der im Kapitel 2 diskutierten Fragestellungen bzgl. der aktuellen Fassung der Norm zur THK (DIN 45681:2005-03 [THK05]) sollen hier näher untersucht werden.

Zum einen wird der Einfluss der maßgeblichen Differenz ΔL (nach Gleichung 2.1) systematisch untersucht. Dabei geht es insbesondere darum, die Erkenntnisse zu den perceptiven Aspekten der THK aus Kapitel 3.3 zu überprüfen.

Eine zentrale Fragestellung ist in diesem Zusammenhang, wo das Maximum der THK-Empfindung liegt und durch welche Art von Signal es hervorgerufen werden kann.

Desweiteren wird die Frequenzabhängigkeit (Kapitel 2.5) der THK analysiert. Damit soll in erster Linie die Annahme der Norm überprüft werden, die Frequenzabhängigkeit der THK sei vollständig durch das Verdeckungsmaß a_v erklärt.

Zur Datenerhebung kommen Paarvergleiche zur Anwendung. Paarvergleiche sind relativ zeitaufwendig, haben aber den großen Vorteil, dass sie für die Probanden einfach durchzuführen sind und selbst kleine Unterschiede zwischen den Stimuli konsistent beurteilt werden können.

5.2. Probanden

An dieser Untersuchung nehmen elf normalhörende, freiwillige Personen aus dem Umfeld des „Graduiertenkolleg Psychoakustik“ sowie der Arbeitsgruppe Akustik des Fachbereichs Physik an der Universität Oldenburg teil.

Alle Teilnehmer haben Erfahrung mit Psychoakustik und THK. Sie werden nicht gesondert für ihre Tätigkeit entlohnt, da sie Angehörige der Universität sind. Zwei Teilnehmer sind weiblich und neun männlich im Alter von 22 bis 30 Jahren (Median 26 Jahre).

Im zweiten Teil der Untersuchung nehmen neun der elf Probanden erneut teil, es entfallen je ein weiblicher und ein männlicher Teilnehmer. Der Median der Altersverteilung dieser Gruppe ändert sich dadurch nicht, ebenso wenig wie das minimale und maximale Alter der Probanden.

5.3. Durchführung

Es wird ein vollständiger Paarvergleich mit jeweils 18 der unten näher beschriebenen Stimuli durchgeführt, d.h. es sind von jedem Probanden 153 einzelne Paarvergleichsbewertungen vorzunehmen.

Die Versuchssteuerung erfolgt mit der Software SI [PK96] auf einer SUN-Workstation. Damit ist es möglich, die einzelnen Paarvergleichsbewertungen in randomisierter Reihenfolge durchzuführen, um Sequenzeffekte zu vermeiden.

Die Darbietung der Stimuli erfolgt diotisch mit dem elektrostatischen Kopfhörer Sennheiser HE60 / HEV70 in einer schallisolierten iac-Abhörkabine.

Die Aufgabe der Probanden besteht darin, die Tonhaltigkeit zu bewerten: „Welches Signal ist *tonhaltiger*?“. Die Urteilsabgabe erfolgt mittels Tastatur der SUN-Workstation durch Drücken der Tasten oder .

Im Vorfeld der Messung werden die Teilnehmer mit der Standard-Orientierung zur Tonhaltigkeit, vgl. Abschnitt 4.7, vertraut gemacht. Die Instruktion der Probanden im Wortlaut findet sich im Anhang E.1.

Die Versuchsdauer für alle 153 Paarvergleichsbewertungen beträgt probandenabhängig etwa 15 bis 20 min.

5.4. Stimuli

Die Stimuli bestehen aus sechs verschiedenen Sinustönen in gleichmäßig anregendem Rauschen (GAR, vgl. Diskussion im Anhang A.1).

Die Frequenzen der Sinustöne betragen 50 Hz, 100 Hz, 470 Hz, 920 Hz, 1450 Hz und 2700 Hz, das entspricht einer Tonheit von 0,5 Bark, 1 Bark, 4,5 Bark, 8 Bark, 11,5 Bark und 15 Bark. Die fünf größten Frequenzen haben somit einen äquidistanten Abstand von 3,5 Bark auf einer gehörgerecht skalierten Frequenzachse. Der Ton mit 50 Hz wird aus den eingangs des Kapitels dargelegten Gründen zusätzlich untersucht.

Die Kombination der sechs Sinustöne mit dem Rauschen erfolgt mit den folgenden Werten für die maßgebliche Differenz ΔL : 0 dB, 12 dB, 24 dB, 36 dB und 48 dB. Zusätzlich wird jeder Ton auch ohne Rauschen untersucht. Diese Kombination wird hier als ∞ dB³² bezeichnet um anzudeuten, dass der reine Ton

³² ∞ wird im Folgenden auch synonym als „Inf“ bezeichnet.

die technisch maximal mögliche maßgebliche Differenz besitzt. Diese Bezeichnung ist symbolisch gemeint, da eine maßgebliche Differenz von ∞ dB praktisch und theoretisch nicht erreicht werden kann. Auf diese Weise ergeben sich 36 Stimuli³³, siehe Tabelle 5.1.

Tabelle 5.1.: Übersicht der Schalle: In den Zeilen die sechs verschiedenen maßgeblichen Differenzen ΔL , in den Spalten die sechs Frequenzen f . Zum Anhören die jeweiligen Schalle anklicken.

$\Delta L \setminus f$	50 Hz	100 Hz	470 Hz	920 Hz	1450 Hz	2700 Hz
Inf	50_Inf	100_Inf	470_Inf	920_Inf	1450_Inf	2700_Inf
48	50_48	100_48	470_48	920_48	1450_48	2700_48
36	50_36	100_36	470_36	920_36	1450_36	2700_36
24	50_24	100_24	470_24	920_24	1450_24	2700_24
12	50_12	100_12	470_12	920_12	1450_12	2700_12
0	50_0	100_0	470_0	920_0	1450_0	2700_0

Die sechs Stimuli mit $\Delta L=0$ dB sind so generiert, dass die Töne sehr nah an der Mithörschwelle liegen, aber bei genauem Hinhören noch wahrnehmbar sind (vgl. dazu Tabelle E.1 im Anhang E).

Die Stimuli mit $\Delta L=\infty$ dB sind reine Töne ohne Rauschen. Für $\Delta L=48$ dB überwiegen die Töne deutlich, das Rauschen ist im Hintergrund aber wahrnehmbar. Die bei den Experimenten mit dem Repertory Grid gefundene verbale Beschreibung für diese Stimulus-Konstellation ist „verrauschter Ton“, da sich die Stimuli so anhören, als wäre das Rauschen technisch bedingt und nicht beabsichtigt.

Alle Schalle werden vor Versuchsbeginn mit einer Abtastrate von 44 100 Hz und einer Amplitudendynamik von 16 bit generiert und gespeichert. Während der Paarvergleiche werden diese Stimuli dann nur noch abgespielt. So wird gewährleistet, dass alle Probanden die exakt gleichen Stimuli bewerten.

Die Länge der Stimuli beträgt 500 ms. Am Anfang und am Ende weisen sie jeweils 20 ms lange lineare Flanken auf, um Knacken zu vermeiden. Die Pause zwischen den beiden im Paarvergleich zu beurteilenden Stimuli beträgt 500 ms.

Der Darbietungspegel wird für alle Stimuli auf 12 sone_{GF} eingestellt. Die Berechnungen dafür erfolgen mit dem Head Acoustics Analysesystem Artemis [Hea98] in einer Version von 1998.

Um die Lautheit von 12 sone_{GF} während der Darbietung praktisch herzustellen und die Stimuli somit zu kalibrieren, reicht es aus, eines der sechs Signale aus

³³ Durch Anklicken der Stimuli in Tabelle 5.1 wird bei geeigneter Konfiguration und verfügbarem Zugriff auf das Internet der jeweilige Schall abgespielt. Alternativ können die Schalle auch direkt von der Webseite www.tonhaltigkeit.de herunter geladen oder angehört werden.

der untersten Zeile der Tabelle 5.1 ($\Delta L=0$ dB) auf 62,3 dB SPL einzustellen. Die angegebene Lautheit von 12 sone_{GF} wird dann für alle Signale erreicht.

5.5. Abhängigkeit der THK von der maßgeblichen Differenz ΔL

Im ersten Teil dieser Untersuchung wird der Einfluss von ΔL untersucht. Dazu werden repräsentativ die Töne mit den Frequenzen $f=100$ Hz, 920 Hz und 2700 Hz ausgewählt und diese bei allen sechs ΔL s getestet, siehe Tabelle 5.2.

Tabelle 5.2.: Schalle zur Untersuchung der maßgeblichen Differenz ΔL (hellblau hinterlegt). In den Zeilen sind die verschiedenen maßgeblichen Differenzen ΔL , in den Spalten die Frequenzen f aufgeführt.

$\Delta L \setminus f$	50 Hz	100 Hz	470 Hz	920 Hz	1450 Hz	2700 Hz
Inf	50_Inf	100_Inf	470_Inf	920_Inf	1450_Inf	2700_Inf
48	50_48	100_48	470_48	920_48	1450_48	2700_48
36	50_36	100_36	470_36	920_36	1450_36	2700_36
24	50_24	100_24	470_24	920_24	1450_24	2700_24
12	50_12	100_12	470_12	920_12	1450_12	2700_12
0	50_0	100_0	470_0	920_0	1450_0	2700_0

Die so gewonnenen Daten zeigen eine sehr gute individuelle Konsistenz für alle elf Teilnehmer, d.h. die Probanden haben die Paarvergleiche anhand eines stabilen, internen Bewertungskriteriums vorgenommen und über die Dauer der Bewertungen weitgehend beibehalten.

Als ein grobes Maß für die individuelle Konsistenz kann die Anzahl der zirkulären Triaden d , i.e. die Anzahl intransitiv bewerteter Paarvergleiche, herangezogen werden [BLB00]. Die Anzahl der zirkulären Triaden liegt für die Teilnehmer zwischen $d=3$ und $d=66$ (Median über alle elf Probanden $d=21$).

Ein normiertes und daher besser geeignetes Maß für die individuelle Konsistenz ist der *individuelle Konsistenzkoeffizient* K . Er ist definiert als Anteil der beobachteten zirkulären Triaden d an den maximal möglichen zirkulären Triaden d_{max} , subtrahiert von 1: $K=1 - d/d_{max}$. K kann daher Werte aus dem Intervall $[0, 1]$ annehmen, wobei "1" für perfekte Konsistenz ohne zirkuläre Triaden steht und "0" keine Konsistenz bedeutet (vergleiche z.B. [BLB00] Seite 489).

Für den durchgeführten Paarvergleich ist die maximal mögliche Anzahl zirkulärer Triaden d_{max} gleich 204. Die individuelle Konsistenz K für die elf Teilnehmer liegt mit diesen Werten zwischen $K=.73$ und $K=.99$, (Median $K=.91$). Selbst für den am wenigsten konsistenten Probanden mit 66 zirkulären Triaden sind die Ergebnisse signifikant auf einem Niveau $p<.001$.

Abbildung 5.1 zeigt die individuellen Daten für alle Probanden. Aufgetragen ist die subjektiv bewertete THK über der maßgeblichen Differenz ΔL . Als Maß für die THK wird die Anzahl gewonnener Paarvergleiche benutzt. Da jeweils 18 Signale verglichen werden, ist die maximal mögliche Anzahl gewonnener Paarvergleiche 17.

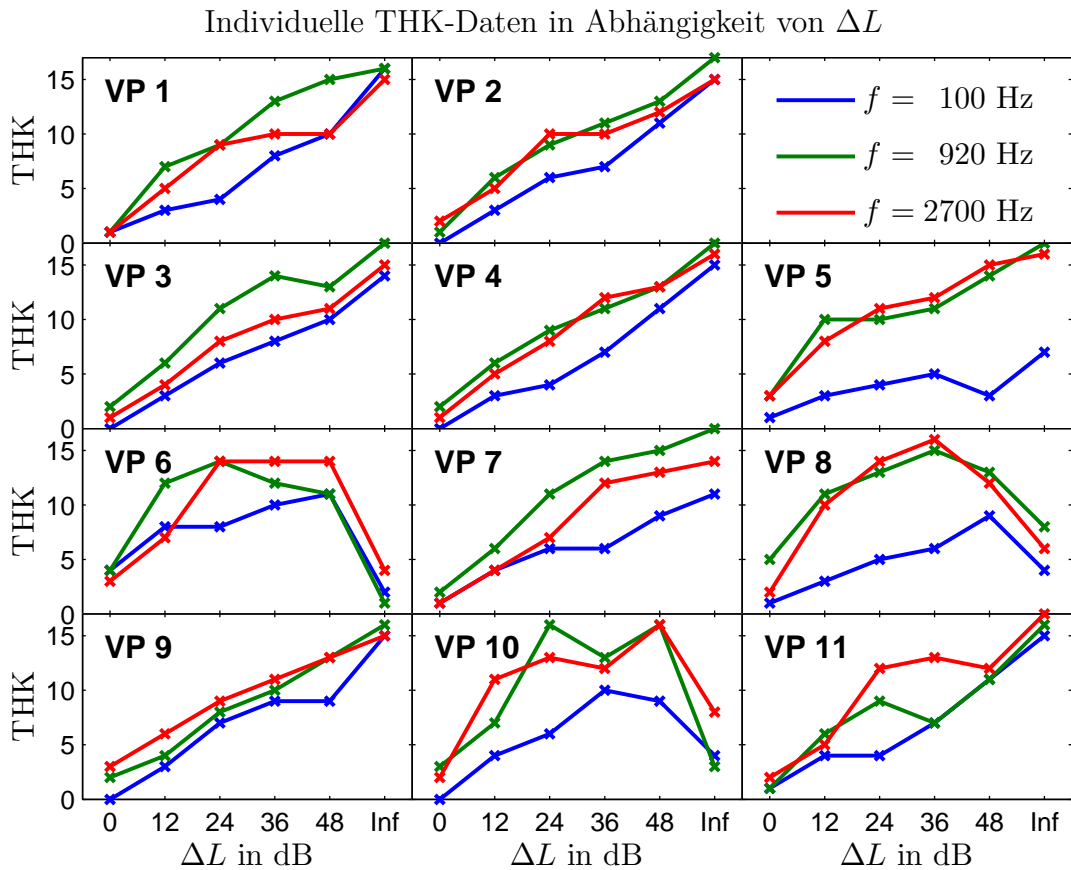


Abbildung 5.1.: Individuelle Daten für die elf Probanden. Aufgetragen ist die Anzahl gewonnener Einzelpaarvergleiche über ΔL für die drei gewählten Frequenzen. Probanden VP 1 bis VP 5, VP 7, VP 9 und VP 11 bilden die Gruppe 1, Probanden VP 6, VP 8 und VP 10 bilden die Gruppe 2.

Acht der elf Teilnehmer bewerten die THK vorrangig anhand der maßgeblichen Differenz ΔL : umso größer die maßgebliche Differenz, desto ausgeprägter wird die wahrgenommene THK empfunden. Diese Probanden werden in Gruppe 1 zusammengefasst. Zu ihnen gehören VP 1 bis VP 5, VP 7, VP 9 und VP 11.

Die anderen drei Teilnehmer (VP 6, VP 8 und VP 10) unterscheiden sich dadurch, dass sie für einen reinen Ton, bzw. ein Signal, das aus einem sehr starken Ton und nur wenig Rauschen („verrauschter Ton“) besteht, ihre Empfindung

5. Tonhaltigkeit von Einzeltönen in Rauschen im vollständigen Paarvergleich

nicht mit dem Begriff THK beschreiben (Gruppe 2). Diese beiden Signale liegen daher außerhalb des Bereiches der THK und werden in diesem Fall stets als weniger tonhaltig bewertet als der beim Paarvergleich alternativ dargebotene Stimulus.

Eine Möglichkeit, die Paarvergleichsdaten auszuwerten und dabei auf intervallskaliertes Skalenniveau zu transformieren, besteht im 1927 von Thurstone beschriebenen *Law of Comparative Judgement* (LCJ, vgl. z.B. Sixtl [Six67]), welches dem in Abschnitt 4.5 beschriebenen *Law of Categorical Judgement* sehr ähnlich ist.

Zunächst wird aus den Paarvergleichsdaten individuell für jeden Probanden die Dominanzmatrix gebildet. Sie enthält für jede im Paarvergleich beurteilte Kombination der Schalle das abgegebene Präferenzurteil. Die individuellen Dominanzmatrizen werden dann zellenweise addiert und die relativen Häufigkeiten gebildet, indem durch die Anzahl der Beurteiler geteilt wird.

Ein robustes Verfahren zur Bestimmung einer Rangreihe der beurteilten Schalle besteht darin, die Zeilensummen der resultierenden Dominanzmatrix zu bilden und nach ihrer Größe auf- oder absteigend anzuordnen. Die so gewonnenen Ergebnisse haben allerdings nur ordinales Skalenniveau, so dass eine quantitative Aussage über die Abstände der Signale untereinander nicht ohne weiteres möglich ist.

Das LCJ [BD06] sieht vor, die relativen Häufigkeiten der Dominanzmatrix zeilenweise zu kumulieren und diese als Flächenanteile der Standardnormalverteilung anzusehen. Im einfachsten Fall werden die Flächenanteile zu z -Werten (d.h. Abszissenwerte der Standardnormalverteilung) transformiert.

Die Anwendung dieser Transformation ist an dieser Stelle jedoch nicht ohne Weiteres möglich, da mehrere Zellen der Dominanzmatrix entweder mit 0 oder 1 belegt sind. Diese Flächenanteile erreicht die Standardnormalverteilung nur asymptotisch und die korrespondierenden Abszissenwerte sind $\pm\infty$, sodass die folgenden Berechnungsschritte nicht durchgeführt werden können.

Abhilfe wird durch solche Transformationen möglich, die endliche Werte für Zahlen x aus dem Bereich $[0:1]$ liefern. Das ist z.B. durch die Verwendung der arcsin-Transformation ($z = \arcsin(2x - 1)$) [DIN89] oder der Z-Transformation [Six67] durch Standardisierung (eigentlich Studentisierung) auf Mittelwert $\mu=0$ und Standardabweichung $\sigma=1$ gegeben: $Z = \frac{x-\mu}{\sigma}$.

Im Folgenden wird die Z-Transformation³⁴ auf die Dominanzmatrix angewendet und aus den so erhaltenen Werten die Zeilenmittelwerte berechnet. Man erhält daraus die intervallskalierten Skalenwerte, indem das Minimum der Zeilensummen zu allen einzelnen Zeilensummen addiert wird. Die so aufbereiteten Daten zeigt Abbildung 5.2 für beide Gruppen.

Als Konkordanzmaß, d.h. als Maß für die Übereinstimmung der Probanden

³⁴ Bei Verwendung der arcsin-Transformation sind die Resultate prinzipiell identisch.

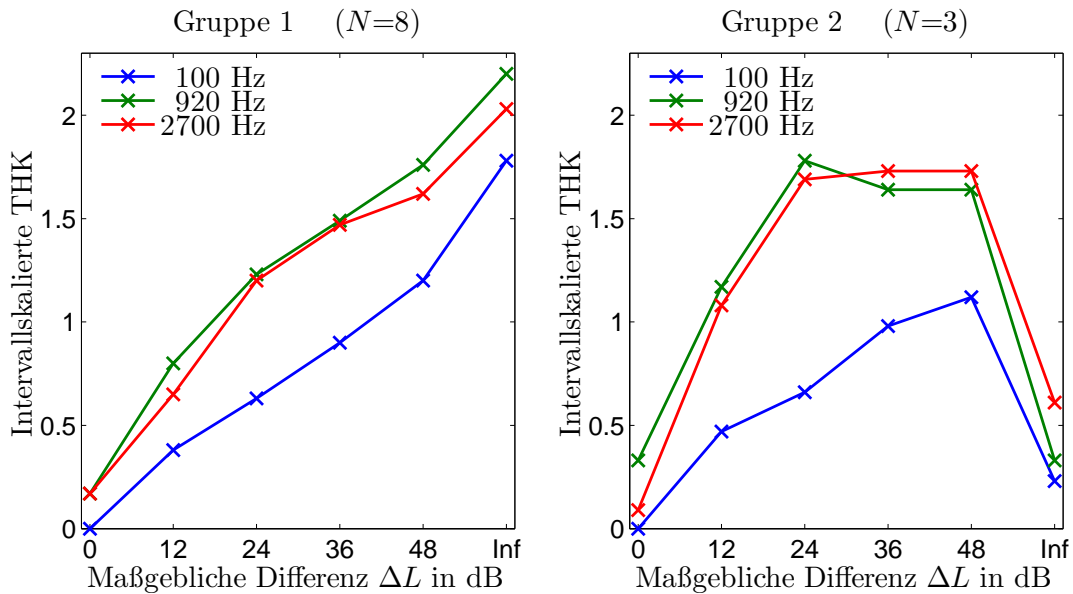


Abbildung 5.2.: Mittels des Law of Comparative Judgement auf Intervallskalenniveau transformierte THK für Gruppe 1 (links) und Gruppe 2 (rechts).

untereinander, wird hier die *Akkordanz* A nach Kendall (vgl. [BLB00] Seite 495) verwendet. Mit der Anzahl der Teilnehmer N gilt für ungerade Werte von N , dass die Akkordanz Werte aus dem Intervall $[-1/(N-1), 1]$ annehmen kann. Ein Akkordanzmaß von $A=1$ deutet auf perfekte Übereinstimmung der Urteile zwischen den Probanden hin. Nimmt das Akkordanzmaß ab, so ist die Übereinstimmung der Probandenurteile entsprechend geringer.

In dieser Untersuchung ist die Akkordanz für Gruppe 1 $A=.70$ auf einem Signifikanzniveau $p<.001$. Dieser Wert ist durchaus veritabel und zeigt, dass die Probanden sehr ähnliche Kriterien für ihre Bewertungen der Paarvergleiche anlegen. Gruppe 2 hat eine etwas geringere, aber immer noch zufrieden stellende Akkordanz von $A=.55$ auf einem Signifikanzniveau $<.001$. Die etwas geringere Akkordanz dieser Gruppe ist damit zu erklären, dass die Probanden die Lage des THK-Maximum leicht verschieden bewerten und die Übereinstimmung daher verringert wird.

In beiden Gruppen ist die Bewertung der THK für die Frequenzen von 920 Hz und 2700 Hz sehr ähnlich, tendenziell sind die Werte für die Frequenz von 920 Hz jedoch leicht größer. Die THK für die Frequenz von 100 Hz wird klar am geringsten eingeschätzt.

Für Gruppe 1 muss die maßgebliche Differenz ΔL für die Frequenz von 100 Hz bis zu etwa 24 dB größer sein als für die beiden anderen Frequenzen, um den gleichen THK-Eindruck hervorzurufen. Gruppe 2 zeigt ein Maximum zwischen

5. Tonhaltigkeit von Einzeltönen in Rauschen im vollständigen Paarvergleich

24 und 48 dB, je nach Frequenz.

Eine ergänzende Randnotiz ist in diesem Zusammenhang, dass die drei Probanden aus Gruppe 2 eine geisteswissenschaftliche Ausbildung haben, während die Probanden aus Gruppe 1 naturwissenschaftlich oder technisch ausgebildet sind. Aufgrund der kleinen Fallzahlen kann hieraus allerdings keine belastbare Aussage gefolgert werden.

Werden die beiden Gruppen trotz der verschiedenen Bewertungsparadigmen der Probanden zusammengefasst, so sinkt die Akkordanz gegenüber der Auswertung in getrennten Gruppen auf $A=.51$ (Signifikanzniveau $p<.001$) ab. Das liegt vorrangig an der unterschiedlichen Bewertung der THK für den „reinen Ton“ mit $\Delta L=Inf$ dB. Wird die Analyse ohne diese Schalle wiederholt, so können die Probanden in einer einzigen Gruppe zusammengefasst werden. Das Ergebnis zeigt Abbildung 5.3.

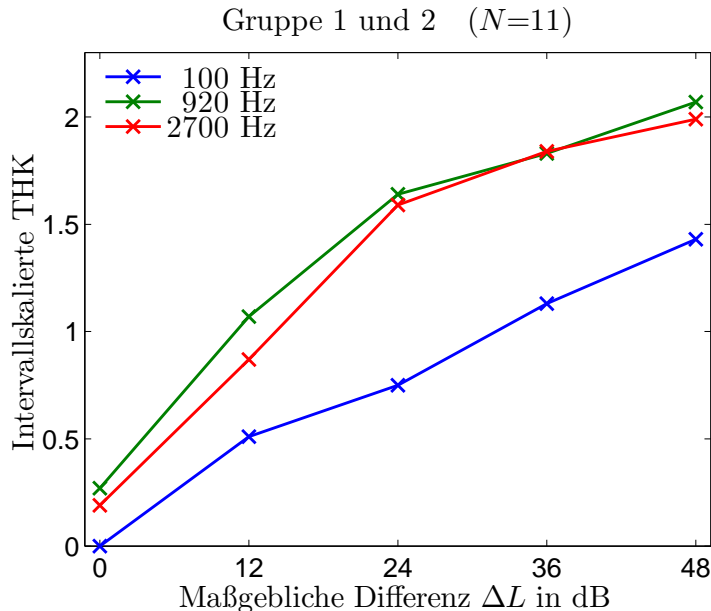


Abbildung 5.3.: Mittels des Law of Comparative Judgement auf Intervallskalenniveau transformierte THK für Gruppe 1 und 2 kumuliert. Die maßgebliche Differenz von Inf dB ist bei dieser Analyse nicht berücksichtigt, siehe Text.

Die Akkordanz für die elf Probanden beträgt in diesem Fall $A=.63$ auf einem Signifikanzniveau $<.001$. Es zeigt sich, dass die Bewertungen der THK monoton mit der maßgeblichen Differenz ΔL ansteigen. Für maßgebliche Differenzen bis etwa 24 dB steigt die THK nahezu linear, darüber hinaus wächst die empfundene THK nur noch wenig an.

5.6. Abhängigkeit der THK von der Frequenz

In diesem Teil wird die Frequenzabhängigkeit der THK evaluiert. Dazu werden die 18 Schalle mit $\Delta L=12$ dB, 36 dB und ∞ dB bei allen verfügbaren Frequenzen bewertet, siehe Tabelle 5.3.

Tabelle 5.3.: Schalle zur Untersuchung der Frequenz (hellgrün hinterlegt). In den Zeilen sind die verschiedenen maßgeblichen Differenzen ΔL , in den Spalten die Frequenzen f aufgeführt.

$\Delta L \setminus f$	50 Hz	100 Hz	470 Hz	920 Hz	1450 Hz	2700 Hz
Inf	50_Inf	100_Inf	470_Inf	920_Inf	1450_Inf	2700_Inf
48	50_48	100_48	470_48	920_48	1450_48	2700_48
36	50_36	100_36	470_36	920_36	1450_36	2700_36
24	50_24	100_24	470_24	920_24	1450_24	2700_24
12	50_12	100_12	470_12	920_12	1450_12	2700_12
0	50_0	100_0	470_0	920_0	1450_0	2700_0

Von den elf Teilnehmern des ersten Teils der Untersuchung nehmen zwei nicht weiter an der Studie teil, so dass neun Teilnehmer verbleiben. Die individuelle Konsistenz für diese neun Teilnehmer ist hoch: sie liegt zwischen $K=.67$ und $K=.94$, (Median $K=.88$) auf einem Signifikanzniveau $<.001$. Die Anzahl der zirkulären Triaden liegt zwischen $d=14$ und $d=80$ (Median $d=28$), die maximal mögliche Anzahl zirkulärer Triaden ist hier wie im ersten Teil $d_{max}=204$. Die individuellen Probandendaten zeigt Abbildung 5.4.

Zunächst werden nur die Probanden analysiert, die im ersten Teil der Untersuchung der Gruppe 1 angehören. Somit werden die Probanden VP 8 und VP 10 an dieser Stelle nicht weiter betrachtet. Von den verbleibenden sieben Probanden empfinden fünf ein Maximum der THK für mittlere Frequenzen. Diese werden als Gruppe 3 zusammen gefasst. Für sie ist die Akkordanz mit $A=.67$ akzeptabel auf einem Signifikanzniveau $<.001$.

In Gruppe 4 werden die Probanden kumuliert, für die die empfundene THK mit steigender Frequenz zunimmt. Das trifft für zwei Probanden zu, die sich mit einer Akkordanz von $A=.70$ auf einem Signifikanzniveau $<.001$ recht ähnlich sind. Die intervallskalierten Daten für die Gruppen 3 und 4 zeigt Abbildung 5.5.

Die Abbildung zeigt für Gruppe 3 ein Maximum zwischen 470 Hz und 1450 Hz für die maßgeblichen Differenzen von $\Delta L=\infty$ dB. Für $\Delta L=12$ dB ist das Maximum bei 920 Hz. Sowohl Gruppe 3 als auch Gruppe 4 zeigt einen Sprung der empfundenen THK für Frequenzen zwischen 100 Hz und 470 Hz. Frequenzen unterhalb von 100 Hz werden deutlich weniger tonhaltig empfunden als Frequenzen

5. Tonhaltigkeit von Einzeltönen in Rauschen im vollständigen Paarvergleich

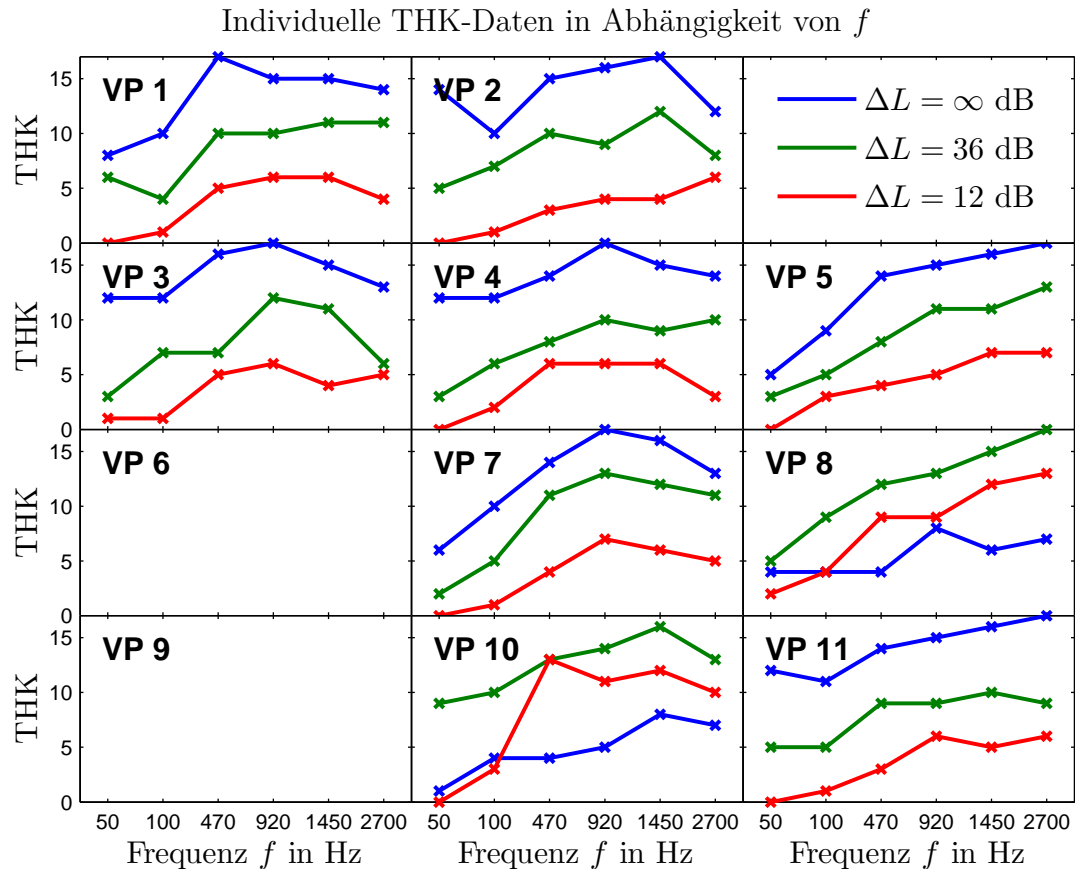


Abbildung 5.4.: Individuelle Daten. Aufgetragen ist die Anzahl gewonnener Einzelpaarvergleiche über f für die drei ausgewählten ΔL .

Probanden VP 6 und VP 9 nehmen nicht mehr teil, die anderen Probanden tragen die selben Bezeichnungen wie in Abbildung 5.1. Probanden VP 1 bis VP 4, VP 7 und VP 10 bilden Gruppe 3. Probanden VP 5, VP 8 und VP 11 gehören Gruppe 4 an.

oberhalb von 100 Hz. Die Vereinfachung der DIN 45681:2005-03 [THK05], Frequenzen unterhalb von 90 Hz nicht mit in die Analyse einzubeziehen, wird durch diese Daten bestätigt. Töne unterhalb von 100 Hz müssen eine vergleichsweise große maßgebliche Differenz aufweisen, um als tonhaltig empfunden zu werden.

Die Ergebnisse des ersten Teils zeigen, dass die reinen Töne eine Sonderstellung einnehmen und nicht ohne Weiteres mit den anderen Signalen verglichen werden können³⁵. Aus diesem Grund werden für die weitere Auswertung alle Paarvergleiche ignoriert, bei denen ein reiner Ton ($\Delta L = \infty$ dB) mit einem beliebigen anderen Signal verglichen wird. Dann kann Proband VP 10 der Gruppe 3

³⁵ Die Akkordanz für alle Probanden der Gruppe 3 inklusive VP 10 beträgt lediglich $A = .46$. Für Gruppe 4 inklusive VP 8 sinkt die Akkordanz auf $A = .29$.

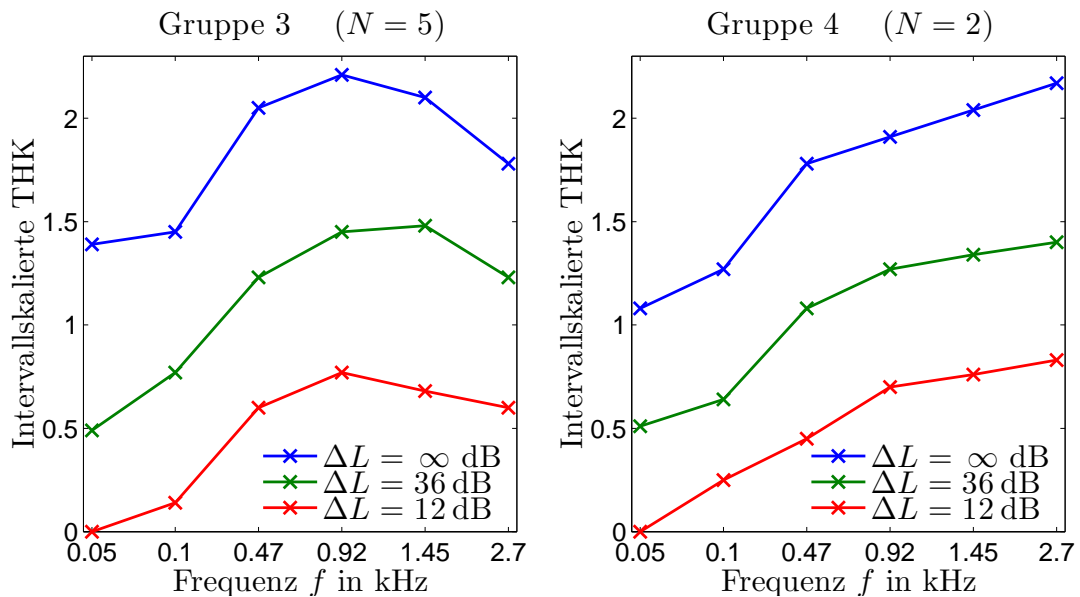


Abbildung 5.5.: Mittels des Law of Comparative Judgement auf Intervallskalenniveau transformierte THK für Gruppe 3 (links) und Gruppe 4 (rechts).

und VP 8 der Gruppe 4 zugeordnet werden. Die Paarvergleichsdaten für beide Gruppen werden erneut per Law of Comparative Judgements auf Intervallskalenniveau transformiert und in Abbildung 5.6 zusammengefasst.

Die Akkordanz für Gruppe 3 aus Abbildung 5.6 beträgt $A=.56$ auf einem Signifikanzniveau $<.001$. Der Verlauf der Tonhaltigkeitsbewertung zeigt für $\Delta L=12$ dB ein wenig ausgeprägtes Maximum bei 920 Hz, die THK wird jedoch für Frequenzen oberhalb 470 Hz im Wesentlichen vergleichbar eingeschätzt. Für die maßgebliche Differenz von $\Delta L=36$ dB ist das Maximum etwas ausgeprägter und liegt bei 1450 Hz. Der weitaus größte Sprung der THK-Bewertung liegt in beiden Fällen zwischen den Frequenzen 100 Hz und 470 Hz, unterhalb von 470 Hz nimmt die Bewertung der THK deutlich ab.

Die Akkordanz für Gruppe 4 aus Abbildung 5.6 beträgt ebenfalls $A=.56$ auf einem Signifikanzniveau $<.001$. Der Verlauf der Tonhaltigkeit steigt annähernd linear mit der Frequenz an.

Der Verlauf der THK für beide Gruppen aus Abbildung 5.6 ist im Wesentlichen vergleichbar zu Abbildung 5.5. Die Akkordanz ist für Gruppe 3 und 4 ohne den reinen Ton etwas niedriger als mit dem reinen Ton, da in diesem Fall nur noch 66 anstelle von 153 Paarvergleichen zur Analyse zur Verfügung stehen. Dafür ist die Anzahl der Probanden in diesem Fall um je einen Teilnehmer größer.

Die DIN 45681:2006-08 unterscheidet nicht zwischen verschiedenen Empfin-

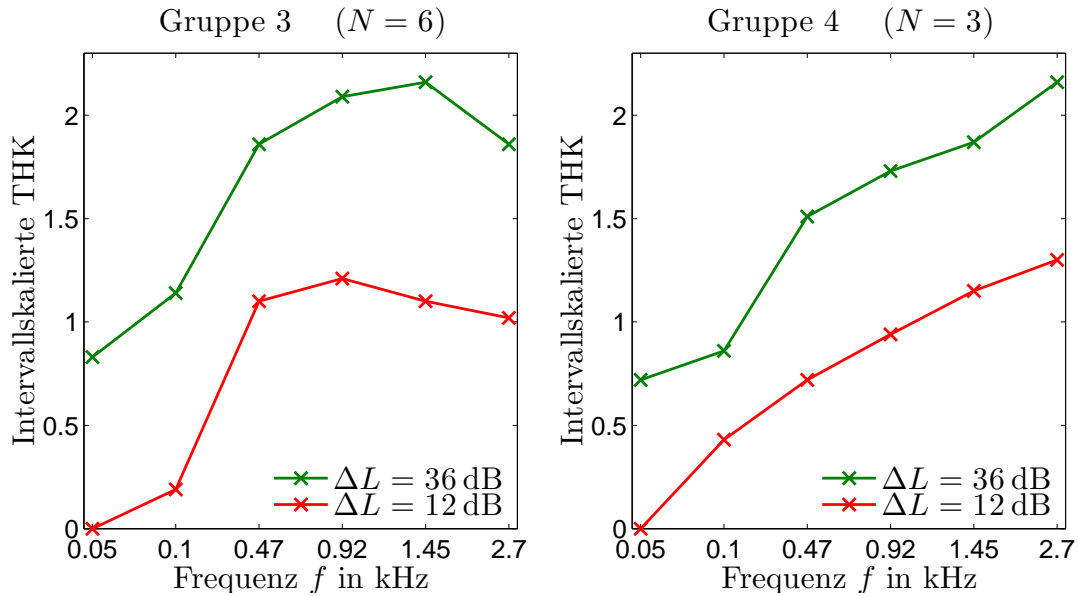


Abbildung 5.6.: Anhand des Law of Comparative Judgement auf Intervallskalenniveau transformierte THK für Gruppe 3 (links) und Gruppe 4 (rechts). Die reinen Töne sind in dieser Auswertung nicht berücksichtigt.

dungen bzgl. der THK, daher werden die beiden Gruppen 3 und 4 abschließend zusammengefasst.

Die Daten für alle Probanden zeigt Abbildung 5.7. Es ist nur die maßgebliche Differenz von 12 dB aufgetragen, da dies dem relevanten Bereich der THK für die gutachterliche Praxis am nächsten kommt. Die Akkordanz für alle Teilnehmer beträgt in diesem Fall erneut $A=.56$.

5.7. Diskussion

Die durchgeführte Untersuchung zeigt, dass die THK angemessen mit Paarvergleichen beurteilt werden kann. Das wird sowohl durch die hohe intra-individuelle Konsistenz als auch die inter-individuelle Konkordanz der Probanden unter Beweis gestellt.

Wie in Übereinstimmung mit der DIN 45681:2006-08 [THK06] zu erwarten, zeigt sich eine wesentlich größere Abhängigkeit der THK von der maßgeblichen Differenz als von der Frequenz. Die prinzipiell unterschiedliche Bewertung der THK zwischen einigen Probanden bei sehr großen maßgeblichen Differenzen ist hingegen ein neuartiger Effekt, der in dieser Form noch nicht berichtet worden ist.

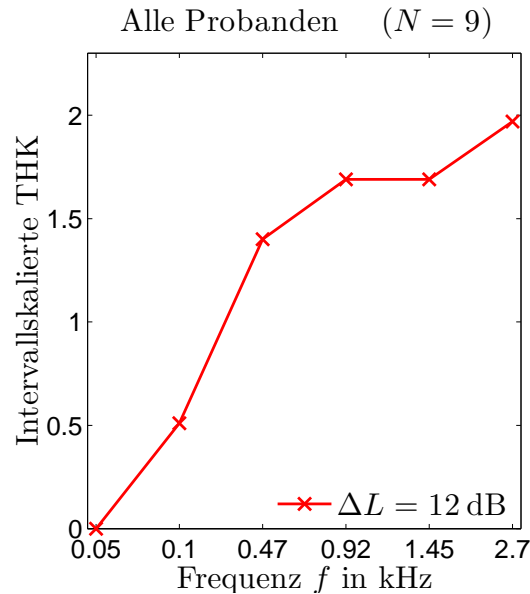


Abbildung 5.7.: Anhand des Law of Comparative Judgement auf Intervallskalenniveau transformierte THK für alle Probanden, d.h. Gruppe 3 und 4 gemeinsam.

Die meisten Beurteiler bewerten die THK in erster Linie anhand der maßgeblichen Differenz ΔL : die subjektiv empfundene THK steigt nahezu linear mit wachsendem ΔL an (Abbildung 5.2, links).

Einige wenige Beurteiler zeigen jedoch ein anderes Verhalten (Abbildung 5.2, rechts). Für ansteigende maßgebliche Differenzen bewerten sie die THK zunächst ebenfalls ansteigend (bis zu etwa 24 dB), ehe sie in einen Bereich maximaler THK über geht (etwa zwischen 24 dB und 48 dB) und dann wieder abfällt (oberhalb von 48 dB). Dieses Verhalten gilt für die untersuchten Frequenzen von 900 Hz und 2700 Hz. Für die deutlich niedrigere Frequenz von 100 Hz tritt das Maximum der THK bei größeren Werten für die maßgebliche Differenz auf (48 dB) und ist etwas schärfer.

Dieses Verhalten ist vergleichbar zu den Ergebnissen aus Kapitel 3.3. Reine Töne, oder Töne bei denen das Rauschen gegenüber den Tönen kaum wahrnehmbar ist, werden nicht als tonhaltig bezeichnet, das Maß THK ist daher für Schalle dieser Art nicht anwendbar.

Der Anwendungsbereich der DIN 45681:2005-03 beschränkt sich im Wesentlichen auf maßgebliche Differenzen unterhalb von 20 dB. Für diesen Bereich sind die diskutierten Bewertungen der THK beider Gruppen vergleichbar. Die unterschiedliche Bewertung der THK für größere maßgebliche Differenzen stellt somit für die DIN 45681:2005-03 kein Problem dar.

5. Tonhaltigkeit von Einzeltönen in Rauschen im vollständigen Paarvergleich

Hinsichtlich der Frequenzabhängigkeit der THK zeigen sich ebenfalls zwei unterschiedliche Beurteilungsstrategien der Beurteiler. Der überwiegende Teil empfindet ein Maximum bei Frequenzen um etwa 1000 Hz. Eine kleine Anzahl von Beurteilern bewertet die THK hingegen monoton ansteigend mit der Tonfrequenz.

Für beide Strategien stellt sich heraus, dass die THK für Frequenzen bis etwa 100 Hz deutlich geringer eingeschätzt wird als darüber. Dies rechtfertigt die Vereinfachung der DIN 45681:2005-03, Frequenzen unterhalb von 90 Hz nicht mit in die Analyse einzubeziehen.

Es zeigt sich weiterhin die Existenz eines, über das frequenzabhängige Verdeckungsmaß $a_v(f)$ hinausgehenden, Effekts der Frequenz auf die THK. Dieser Effekt ist nicht zu erwarten, da das Verdeckungsmaß $a_v(f)$ über die maßgebliche Differenz ΔL bei der Erstellung der Stimuli berücksichtigt worden ist. Aus diesem Grund sollten die Stimuli jeder Zeile aus Tabelle 5.1 nach DIN 45681:2005-03 jeweils die gleiche THK hervorrufen.

Da bzgl. der THK somit keine Unterschiede für diese Stimuli vorliegen, sollten sie im Paarvergleich in zufälliger Weise bewertet werden. Durch diese zufälligen Bewertungen würden folglich sehr viele zirkulären Triaden hervorgerufen.

Die Anzahl der zirkulären Triaden ist jedoch verhältnismäßig klein (siehe Abschnitt 5.6), was die Annahme von zufälligen Bewertungen widerlegt. Es existiert also eine Abhängigkeit der THK von der Frequenz, die über die in der DIN 45681:2005-03 berücksichtigte Frequenzabhängigkeit hinaus geht.

Aus diesem Grund wird die Untersuchung zur Frequenzabhängigkeit der THK in Abschnitt 6.4.8 mit einer neuen Methode vertieft. Dort sind ebenfalls weitere Diskussionen zu diesem Themenkomplex aufgeführt.

6. Einfluss verschiedener Signalparameter auf die Tonhaltigkeit

6.1. Motivation und Ziel der Untersuchung

Wie schon in Kapitel 2 diskutiert, kann bzw. darf die Norm zur THK DIN 45681: 2006-08 [THK06] nicht auf beliebige Schalle angewendet werden. Ein Beispiel, bei dem die Norm versagt, sind Schalle mit „heulenden“ oder „kreischenden“ Tönen.

Diese komplexen Töne rufen zwar eine THK hervor, die Norm kann aber mangels ausreichender physikalischer Kriterien zur Beschreibung der Töne nicht angewendet werden. Damit ist eine objektive Bewertung der THK anhand der DIN 45681 nicht möglich.

In der Regel ist in solchen Fällen die ausschließlich subjektive Beurteilung (durch einen Gutachter) die einzige Möglichkeit, die THK dieser Schalle trotzdem zu erfassen.

Diese Einschränkungen bei der Anwendung der Norm können mit dem im folgenden beschriebenen Verfahren vermieden bzw. umgangen werden. Auf diese Weise kann die THK für beliebige Schalle objektiviert werden. Die Idee dabei ist, absolute (Gutachter-) Urteile, die im wesentlichen von der internen Referenz des Beurteilers abhängen, durch ein objektives Verfahren auf eine vorgegebene, externe Referenz zu verlagern. Die THK für diese externe Referenz kann im Weiteren mit dem in der Norm vorgeschlagenen Verfahren berechnet und als Maß für das zu bewertende Schallsignal betrachtet werden.

Das Verfahren funktioniert so, dass ein bezüglich der THK einstellbares *Referenzsignal* auf die gleiche THK wie ein beliebiger zu bewertender Schall eingeregelt wird. Eine sehr ähnliche Verfahrensweise ist von der *Pegellautstärke*³⁶ bekannt: Die (empfundene) Lautstärke eines beliebigen Signals wird indirekt skaliert, indem ein Referenzton mit einer Frequenz von 1000 Hz subjektiv auf die gleiche Lautstärke eingestellt wird wie das zu bewertende Signal. Der Schalldruckpegel des Referenztones wird dann als Maß für die Lautstärke des zu bewertenden Signals angegeben.

³⁶ Die Pegellautstärke wird aus historischen Gründen auch als *Lautstärkepegel* bezeichnet.

Das so definierte Verfahren für die Pegellautstärke besitzt in der Psychoakustik eine sehr große Bedeutung, weshalb dafür eine eigene Einheit, das *phon*, eingeführt worden ist. Die Pegellautstärke in phon ist gleich dem Schalldruckpegel des Referenztones in dB.

Die Pegellautstärke konnte sich deshalb durchsetzen, weil es für die Messung der empfundenen Lautstärke lediglich notwendig ist, einen direkten Vergleich durchzuführen. Ein Vergleich ist eine relativ einfache und robuste Methode der Skalierung, da er keine Vorkenntnisse erfordert und sofort ausgeführt werden kann. Aus diesem Grunde wird angenommen, dass ein solches Verfahren auch für die Bewertung der THK zielführend einsetzbar ist.

Die Intention des zu entwickelnden Verfahrens besteht darin, jeden beliebigen Schall bzgl. der THK bewerten zu können, indem die empfundene THK eines geeigneten Referenzsignals gleich eingestellt wird. Die THK des Referenzsignals wird dann als Maß für die THK des zu bewertenden Schalls angesehen.

Somit wird es möglich, z.B. Schalle mit „heulenden“ oder „kreischenden“ Tönen deutlich objektiver zu beurteilen als bisher. Auch zeitlich fluktuierende Schalle werden mit diesem Verfahren bewertbar.

Bei zweckmäßiger Wahl des Referenzsignals ist es daher immer realisierbar, die THK für das Referenzsignal direkt mit der aktuellen Norm DIN 45681:2006-08 zu berechnen. Bei geeigneter Normierung des Referenzsignals ist es auch möglich, die THK direkt als Ergebnis des Vergleichsprozesses abzulesen.

Im folgenden Kapitel wird daher zunächst das Verfahren für die THK entwickelt bzw. beschrieben und die Anwendbarkeit im Folgenden untersucht. Dazu werden einige wichtige Parameter, von denen angenommen wird sie könnten einen Einfluss auf die THK haben, in mehreren explorativen Untersuchungen studiert.

6.2. Adaptives THK-Messverfahren

6.2.1. Verfahren

Das Referenzsignal muss eindeutig, exakt und reproduzierbar auf die gleiche THK wie der zu analysierende Schall eingestellt werden. Daher ist es entscheidend, ein geeignetes Verfahren für die systematische Einstellung des Referenzsignals auszuwählen. Aus diesem Grund werden folgende Anforderungen an das Einstellverfahren festgelegt:

- Das Verfahren muss eindeutige und reproduzierbare Ergebnisse liefern.
- Das Verfahren soll universell einsetzbar sein, d.h. für möglichst alle vorstellbaren Geräusche bzw. Schalle gleichermaßen angewendet werden können.

nen. Die Minimalanforderung ist eine Anwendbarkeit mindestens auf die technisch relevanten Schalle.

- Das Verfahren soll möglichst einfach sein. Das gilt sowohl bzgl. der Durchführbarkeit für die Probanden (Beurteiler) als auch für die Anwendung, Auswertung und Interpretation der Ergebnisse.
- Zudem ist eine hohe Effektivität wünschenswert. Das heißt, die Resultate können mit geringem Aufwand und in kurzer Zeit erzielt werden.

Das 2AFC-Verfahren erfüllt diese Anforderungen. Es ist wohldefiniert und in vielen psychoakustischen Untersuchungen für ähnliche Zwecke erfolgreich verwendet worden. Daher bietet es sich an, für die Einstellung des Referenztones auf die gleich empfundene THK wie das zu bewertende Signal verwendet zu werden.

Das 2AFC-Verfahren funktioniert so, dass der Beurteiler sich für eine von zwei Antwortalternativen entscheiden muss – er wird demnach immer zu einer Aussage gezwungen.

In dieser Studie wird dem Beurteiler das THK-Referenzsignal und der zu bewertende Schall akustisch dargeboten. Die Aufgabe des Beurteilers besteht darin, zu entscheiden, welches der beiden Signale tonhaltiger ist. Die hier praktizierte Kopplung der Alternativen an zwei verschiedene Signalpräsentationen („Intervalle“) wird auch als 2-Intervall 2AFC-Verfahren (2I-2AFC) bezeichnet.

Vertiefende Informationen bzgl. des AFC-Verfahrens finden sich z.B. in Hartmann [Har05] oder Gelfand [Gel98].

Um die Einstellung der Signale auf gleiche THK zu realisieren, wird eine adaptive Pegelsteuerung (die direkt die THK des Referenzsignals verändert) durchgeführt: in Abhängigkeit von den gegebenen Antworten des Beurteilers wird das Referenzsignal bzgl. der THK so verändert, dass es die gleiche THK hervorruft wie der zu bewertende Schall. Dafür können verschiedene Methoden zur Pegelsteuerung angewendet werden. Die für die Praxis relevantesten Methoden werden von Levitt [Lev71] beschrieben.

Für das adaptive THK-Messverfahren wird ein *einfaches Up-Down Verfahren*³⁷ durchgeführt. Bei dieser Methode werden zwei Stimuli nacheinander in zufälliger Reihenfolge präsentiert. Einer der beiden Stimuli ist der *Teststimulus*³⁸, dessen THK bewertet werden soll. Der andere Stimulus ist der *Vergleichsstimulus*³⁸, der adaptiv verändert wird.

Im Prinzip werden auf diese Weise nacheinander einfache Paarvergleiche („Trials“) für sich ändernde Signale durchgeführt. Eine schematisierte Darstellung eines typischen Messablaufes mit diesem Verfahren zeigt Abbildung 6.1.

³⁷ Das einfache Up-Down Verfahren ist auch als *Staircase* oder *1-up 1-down Methode* bekannt.

³⁸ Nomenklatur nach Hellbrück [HEKZ08] .

Wird die THK des Vergleichsstimulus kleiner beurteilt als die des Teststimulus, so wird jeweils im nächsten Trial die THK des Vergleichsstimulus erhöht, andernfalls verringert. Auf diese Weise wird der Vergleichsstimulus adaptiv so lange verändert, bis seine THK hinreichend genau der des Teststimulus entspricht.

Der so gewonnene Messwert wird als *point of subjective equality* (PSE³⁹) bezeichnet, da die THK von Test- und Vergleichsstimulus für diesen Fall übereinstimmt. In Abbildung 6.1 ist das Ergebnis der THK-Evaluation, d.h. der PSE, durch die grüne horizontale Linie gekennzeichnet.

Die Änderung der THK des Vergleichsstimulus erfolgt zu Beginn mit einer großen *Anfangsschrittweite*, so dass die THK des Vergleichsstimulus zunächst grob gegen die THK des Teststimulus konvergiert, siehe Abbildung 6.1.

Im Verlauf der Messung wird die *Schrittweite* der adaptiven Änderung für die THK des Vergleichsstimulus sukzessive verringert (typischerweise halbiert), in der schematischen Darstellung von acht über vier und zwei Einheiten bis hin zur *minimalen Schrittweite* S_{min} , die in Abbildung 6.1 eine Einheit beträgt.

Die Halbierung erfolgt immer an einem *oberen Umkehrpunkt*, d.h. wenn die THK im voran gegangenen Trial erhöht worden ist und nun aufgrund des aktuellen Vergleiches erniedrigt wird. Das gilt z.B. für Trial Nr. 6, 10, 15, 21, usw.

Sobald die minimale Schrittweite erreicht ist, beginnt die *Messphase*. Aus den ab jetzt gemessenen Umkehrpunkten (obere und untere) wird das Endergebnis (PSE) auf geeignete Weise berechnet.

Für das adaptive THK-Messverfahren wird der Median der Umkehrpunkte gebildet und als Ergebnis des Verfahrens betrachtet, siehe Abbildung 6.1. Als Streuungsparameter und somit als Gütemass, wird der Interquartilbereich ermittelt.

Ein Vorteil des 2I-2AFC Verfahrens besteht darin, dass die Messung durch Änderung der o.g. Parameter leicht konfigurierbar ist. So kann, je nach Anforderung, die Messung in einem gewissen Rahmen schneller oder genauer durchgeführt werden.

6.2.2. THK Referenzsignal

Für das oben beschriebene Verfahren muss der Vergleichsstimulus geeignet ausgewählt und realisiert werden. Da er im Weiteren als *THK-Referenz* dient, ist

³⁹ Der PSE ist der 50%-Wert der psychometrischen Funktion, vgl. auch Literatur zur Signalentdeckungstheorie, z.B. Hellbrück [Hel93], Zimbardo und Gerrig [ZG04] oder Levitt [Lev71]. Für diesen Punkt bewertet ein Proband in 50% der Fälle den Vergleichsstimulus, in den verbleibenden 50% den Teststimulus als tonhaltiger.

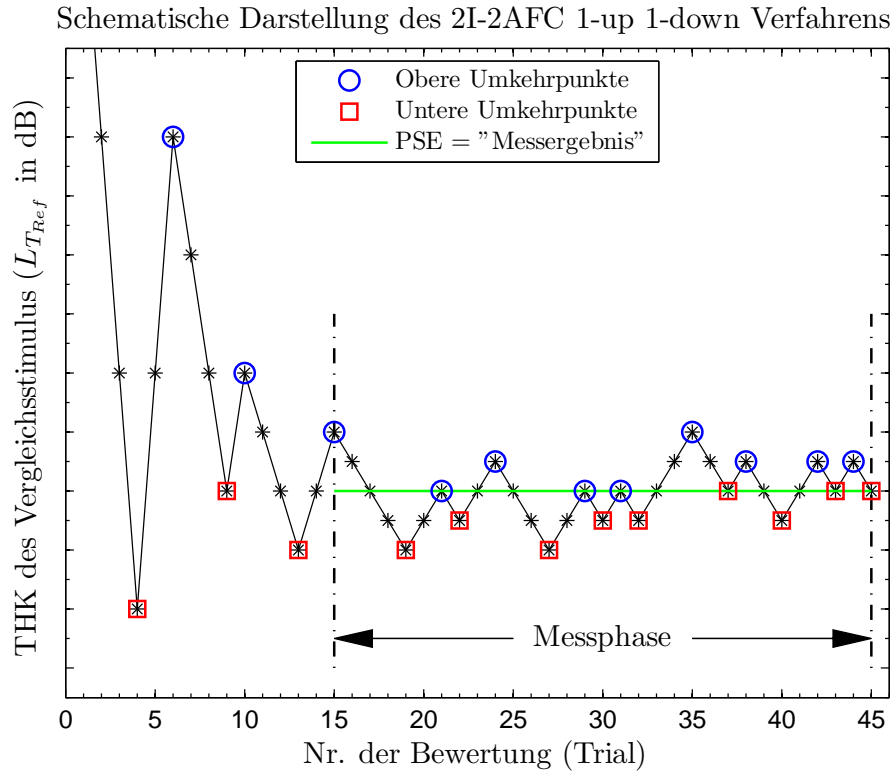


Abbildung 6.1.: Schematische Darstellung des Verlaufes einer 2I-2AFC 1-up 1-down Messung. Aufgetragen ist der Pegel des Vergleichsstimulus $L_{T_{Ref}}$ als Maß für die THK über die nacheinander durchgeführten Paarvergleiche („Trials“) in dB bei einem willkürlichen Bezugspegel. Die Schrittweite wird sukzessive von acht über vier und zwei auf ein dB reduziert. Erst danach beginnt die Messphase. Die grüne Linie symbolisiert den PSE, d.h. das Ergebnis der Messung. Die oberen und unteren Umkehrpunkte sind zur Verdeutlichung farblich markiert.

er ein entscheidender Faktor für die Anwendbarkeit des gesamten Verfahrens.

Das Augenmerk liegt darauf, die THK-Referenz (d.h. den Vergleichsstimulus) universell einsetzbar, wohl definiert und repräsentativ für das Phänomen THK zu erzeugen.

Das Referenzsignal für die THK muss nach den Erkenntnissen aus Kapitel 3.4 aus einem Rauschen (Hintergrund) und einer tonartigen Komponente (Vordergrund) bestehen. Beide Signalkomponenten werden im Folgenden getrennt diskutiert.

Hintergrund des THK Referenzsignals

Der Hintergrund wird so gewählt, dass er für die meisten Fälle, in denen die DIN 45681:2006-08 eingesetzt werden soll, anwendbar ist.

Die DIN 45681 ist insbesondere relevant für Industrie- und Umweltgeräusche. Es ist daher naheliegend, den Hintergrund so zu wählen, wie er bei diesen Schallereignissen meistens auftritt.

Einige typische tonhaltige Geräusche aus der gutachterlichen Praxis wurden von Pompetzki [Pom98] zusammengestellt, siehe Anhang B.1. Sie haben gemeinsam, dass ihr Hintergrund überwiegend rauschhafter Natur ist, allerdings unterscheiden sich die Spektren der Signale beträchtlich.

Im Einzelfall beinhalten sie weitere psychoakustische Phänomene, sind moduliert oder ändern sich auf andere Weise über die Zeit. Das macht diesen Typ Hintergrundgeräusch für die hier angestrebte Verwendung ungeeignet, obwohl die Geräusche an sich ökologisch valide sind.

Stattdessen wird ein exakt zu beschreibendes und signaltechnisch leicht herzustellendes Rauschen verwendet. Zusätzlich wird das Rauschen der Einfachheit halber zeitlich stationär⁴⁰ gewählt.

Dafür kommen prinzipiell die im Anhang A.1 diskutierten Rauschtypen in Frage. Die ersten drei Rauscharten (weiß, rosa, braun) sind eher technisch motiviert, während die beiden folgenden aus psychoakustischer Sichtweise geeigneter erscheinen. Das gleichmäßig anregende Rauschen (GAR) ist deshalb angemessen, weil in jeden auditorischen Filter die gleiche Energie fällt und somit der Frequenzgruppenpegel für alle Frequenzen konstant ist. Bei bekanntem bzw. gemessenem Pegel der tonalen Komponente ergibt sich die maßgebliche Differenz nach DIN 45681:1992-01e [THK92] unmittelbar aus der Differenz des Frequenzgruppenpegels zum Pegel der tonalen Komponente und dem konstant angenommenen Verdeckungsmaß.

Bei Verwendung von gleichmäßig verdeckendem Rauschen (GVR) wird die Ermittlung der THK nach DIN 45681:2006-08 [THK06] besonders einfach, da die tatsächliche (frequenzabhängige) Mithörschwelle durch das Verdeckungsmaß berücksichtigt wird. Die maßgebliche Differenz ist somit identisch zur Differenz zwischen dem Frequenzgruppenpegel und dem Pegel der tonalen Komponente.

Zum Zeitpunkt der Durchführung dieser Studie war die DIN 45681:2006-08 noch nicht veröffentlicht. Daher wird auf die Norm DIN 45681:1992-01e Bezug genommen und somit GAR als Hintergrund ausgewählt. Es ergeben sich dadurch einige Vorteile, die im Folgenden diskutiert werden.

Bei gegebenem Gesamtpegel des GAR L_{GAR} hängt der Frequenzgruppenpegel

⁴⁰ Ein Signal wird in Anlehnung an [OSB04] als stationär bezeichnet, wenn sowohl sein Spektrum als auch sein Pegel im Mittel gleich bleiben. Für die hier diskutierten Rauschen sind quasistationäre Signale ausreichend, d.h. Signale, die nur innerhalb eines bestimmten Zeitintervalls stationär sind.

$L_{G_{GAR}}$ ausschließlich vom Frequenzbereich ab, in dem das Rauschen dargeboten wird. Wird dieser Bereich durch die Anzahl der angeregten Frequenzgruppen N_{FG} angegeben, ergibt sich:

$$(6.1) \quad L_{G_{GAR}} = L_{GAR} - 10 \log(N_{FG}) \text{ dB.}$$

Das so berechnete $L_{G_{GAR}}$ ist für alle Frequenzgruppen unabhängig von der Mittenfrequenz gültig.

In der vorliegenden Studie erstreckt sich das Rauschen immer über den kompletten, bei der Schalldarbietung verwendeten Frequenzbereich. Das heißt die untere Darbietungsfrequenz ist konstant kleiner als 20 Hz, die höchste Frequenz hängt ausschließlich von der Nyquist-Frequenz, d.h. der halben Abtastrate $f_S/2$, ab.

Die Anzahl angeregter Frequenzgruppen ergibt sich, indem die Nyquist-Frequenz in die Tonheit z umgerechnet wird, zu:

$$(6.2) \quad N_{FG} \Rightarrow z_{max} = z \left(\frac{f_S}{2} \right).$$

Die Anzahl der angeregten Frequenzgruppen N_{FG} kann entweder direkt aus Zwickers Originalarbeit zu den Frequenzgruppen [Zwi61] abgelesen oder mittels folgender Näherungsformel nach Zwicker und Terhardt [ZT80] mit $f = f_S/2$ berechnet werden:

$$(6.3) \quad z = 13 \arctan \left(0,00076 \frac{f}{\text{Hz}} \right) + 3,5 \arctan \left(\left(\frac{f}{7500 \text{ Hz}} \right)^2 \right).$$

Tonale Komponente des THK Referenzsignals

Entsprechend der DIN 45681:2006-08 [THK06] (Abschnitt 5.3.4) ist ein Sinuston dasjenige Signal, welches die ausgeprägteste Tonhöhe und damit die größte THK hervorruft. Außerdem ist ein Sinuston ein analytisch sehr einfach zu beschreibendes Signal.

Eine allgemeinere Betrachtung bzgl. der Ausgeprägtheit von Tonhöhen liefern Fastl und Zwicker [FZ07], Kapitel 5.7. Die für die THK relevanten Signale, die eine Tonhöhe hervorrufen, sind: reine Töne⁴¹, (harmonisch) komplexe Töne, Schmalbandrauschen und amplitudenmodulierte Töne. Die Ausgeprägtheit der Tonhöhe für diese Signale nimmt in der Reihenfolge, in der sie aufgeführt sind, ab.

Bei Normierung der Ausgeprägtheit der hervorgerufenen Tonhöhe auf die Ausgeprägtheit eines reinen Tones mit gleicher Tonhöhe zeigen sich folgende

⁴¹ Die Bezeichnungen „reine Töne“ und „Sinustöne“ werden hier synonym verwendet.

Abstufungen: (harmonisch) komplexe Töne haben etwa 50% bis über 80% der Ausgeprägtheit des reinen Tones, Schmalbandrauschen liegen je nach Beschaffenheit etwa im gleichen Bereich und amplitudenmodulierte Töne bewirken eine Ausgeprägtheit zwischen 40% und 60%, bezogen auf den reinen Ton.

Darüber hinaus existieren zahlreiche weitere Signale, die Tonhöhen hervorrufen. Diese sind jedoch deutlich weniger ausgeprägt als die zuvor Diskutierten oder für die THK im Sinne der Norm DIN 45681:2006-08 nicht relevant.

Weiterhin zeigt sich auch an den Ergebnissen aus Kapitel 3, dass es günstig ist einen reinen Ton als THK-Referenz zu verwenden, besonders wenn „Laien“ die THK bewerten sollen. Auf die gleiche Weise verfährt zum Beispiel auch Hansen [Han10].

Desweiteren ist es testmethodisch von Vorteil, das Signal mit der ausgeprägtesten Tonhöhe zur Skalierung zu benutzen. Auf diese Weise erfolgen die Urteile erwartungsgemäß ausschließlich in eine Richtung: die Ausprägung der Tonhöhe für jedes andere Signal als den reinen Ton muss zwangsläufig kleiner bewertet werden.

Der Ton wird, analog zum Rauschen, stationär generiert. Dieses Vorgehen ist der DIN 45681:2006-08 am angemessensten, da sie nur begrenzt für zeitlich fluktuierende Signale anwendbar ist. Weiterhin bleibt das Verfahren damit so einfach wie möglich.

Als freier Parameter für den Referenzton bleibt somit nur noch die Frequenz des Tones $f_{T_{Ref}}$. Diese wird in den folgenden Untersuchungen als $f_{T_{Ref}}=700$ Hz definiert. Die Wahl ist durch nachfolgend aufgeführte Studien motiviert.

Die Frequenz von 700 Hz befindet sich zum einen ungefähr mittig in dem Frequenzbereich, der für die THK von typischen Umweltschallen relevant ist, siehe dazu die Umweltschalle von Pompetzki [Pom98] im Anhang B.1.

Zum anderen zeigen Terhardt et al. [TSS82] in einer Untersuchung zur Ausgeprägtheit von Tonhöhen, dass bei 700 Hz das Maximum für die frequenzabhängige Wichtung des Beitrags zur Ausgeprägtheit der Tonhöhe vorliegt⁴².

Desweiteren liegt die gewählte Frequenz im Bereich „spektraler Dominanz“ nach Ritsma [Rit67].

Eine Untersuchung von Fletcher und Galt [FG50] zeigt, dass der „Beitrag einzelner Frequenzbereiche für die Sprachverständlichkeit“ ein Maximum bei 700 Hz⁴³ besitzt. De facto sind die beiden Kurven von Terhardt et al. sowie Fletcher und Galt nahezu proportional zueinander [TSS82].

In der Musik hat die Frequenz von 700 Hz ebenfalls eine tragende Bedeutung. Sie entspricht annähernd der Note f⁴⁴, bzw. F5⁴⁴, die auf einer Klaviatur auf der

⁴² Terhardt et al. [TSS82], Abbildung 3 auf Seite 683: der frequenzabhängige Beitrag hat einen Verlauf ähnlich einer „Glockenkurve“ mit einem Maximum bei 700 Hz. Dieser Zusammenhang wird auch von Vormann [Vor95] diskutiert.

⁴³ [FG50], Abbildung 5 auf Seite 94

⁴⁴ Die Note f⁴⁴ bzw. F5 entspricht exakt einer Frequenz von 698,456 Hz.

59. von 88 Tasten liegt und ein wesentlicher Bestandteil westlicher Musik ist.

Die aufgeführten Studien zeigen die Relevanz der Frequenz 700 Hz in unterschiedlichen Bereichen wie Tonhöhenwahrnehmung, Sprachwahrnehmung oder Musik. Es ist daher ein erfolgversprechender Ansatz, die Frequenz des Referenztones zunächst als 700 Hz anzunehmen.

Berechnung der THK für das Referenzsignal

Ein Maß für die THK des Referenzsignals ist die maßgebliche Differenz ΔL des Tonpegels $L_{T_{Ref}}$ zum Pegel der Mithörschwelle für diesen Ton ($f_{T_{Ref}}=700$ Hz).

Die Mithörschwelle wird in der zurzeit gültigen Norm DIN 45681:2006-08 nach Gleichung 2.1 berechnet aus dem Frequenzgruppenpegel des verdeckenden Geräusches (in diesem Fall $L_{G_{GAR}}$) um die Tonfrequenz ($f_{T_{Ref}}$) und dem frequenzabhängigen Verdeckungsmaß für den Referenzton $a_v(f_{T_{Ref}})$. Daraus ergibt sich die maßgebliche Differenz für das Referenzsignal zu

$$(6.4) \quad \Delta L_{Ref} = L_{T_{Ref}} - L_{G_{GAR}} - a_v(f_{T_{Ref}}).$$

Nach Gleichung A.1 wird das Verdeckungsmaß für die DIN 45681:2006-08 angenommen als $a_v(f_{T_{Ref}})=-2,5$ dB:

$$(6.5) \quad \Delta L_{Ref} = L_{T_{Ref}} - L_{G_{GAR}} + 2,5 \text{ dB}.$$

Im Normen-Vorschlag von 1992 (DIN 45681E:1992-01 [THK92]) wird das Verdeckungsmaß unabhängig von der Frequenz zu $a_{v92}=-6$ dB angenommen. Die maßgebliche Differenz für diese Version der Norm ΔL_{Ref92} ist somit

$$(6.6) \quad \Delta L_{Ref92} = L_{T_{Ref}} - L_{G_{GAR}} + 6 \text{ dB}.$$

Für das Referenzsignal hängt die THK, d.h. die maßgebliche Differenz ΔL_{Ref} bzw. ΔL_{Ref92} , und somit der Tonzuschlag K_T bzw. K_{T92} ausschließlich vom während der Messung adaptiv eingestellten Tonpegel $L_{T_{Ref}}$ ab, da der Frequenzgruppenpegel für das GAR $L_{G_{GAR}}$ eine vorgegebene Konstante darstellt. Das vorgestellte Verfahren zur THK-Bewertung liefert somit direkt ein anschaulich interpretierbares Maß für die THK, siehe auch den Abschnitt zur grafischen Darstellung der Ergebnisse (6.3.3).

Der für das diskutierte Referenzsignal gültige Tonzuschlag ist zur besseren Übersicht in Tabelle 6.1 angegeben - der Tonzuschlag kann ebenso aus den Tabellen der entsprechenden Normen direkt entnommen werden.

Der frequenzabhängige Unterschied $\delta L(f)$ zwischen den beiden Normvorschlägen zur Berechnung der maßgeblichen Differenz wird definiert als

$$(6.7) \quad \delta L(f) := \Delta L_{Ref92} - \Delta L_{Ref} = 6 \text{ dB} + a_v(f_{T_{Ref}}).$$

Für das hier verwendete Referenzsignal ergibt sich $\delta L(f_{T_{Ref}})$ damit zu 3,5 dB. Der Tonzuschlag K_{T92} kann nach Berücksichtigung dieser Korrektur aus dem rechten Teil der Tabelle 6.1 entnommen werden.

6. Einfluss verschiedener Signalparameter auf die Tonhaltigkeit

Tabelle 6.1.: Tonzuschlag K_T für Messungen mit dem vorgeschlagenen *THK-Referenzsignal*. Die linken Spalten haben Gültigkeit für DIN 45681:2006-08 [THK06]. Die rechten Spalten gelten analog für DIN 45681E:1992-01 [THK92]. Für den Referenzton mit 700 Hz ist $\delta L=3,5$ dB.

ΔL_{Ref} in dB	K_T in dB	$\Delta L_{Ref_{92}}$ in dB	$K_{T_{92}}$ in dB
$\Delta L_{Ref} \leq 0$	0	$\Delta L_{Ref} + \delta L \leq 0$	0
$0 < \Delta L_{Ref} \leq 2$	1	$0 < \Delta L_{Ref} + \delta L \leq 2$	1
$2 < \Delta L_{Ref} \leq 4$	2	$2 < \Delta L_{Ref} + \delta L \leq 4$	2
$4 < \Delta L_{Ref} \leq 6$	3	$4 < \Delta L_{Ref} + \delta L \leq 6$	3
$6 < \Delta L_{Ref} \leq 9$	4	$6 < \Delta L_{Ref} + \delta L \leq 8$	4
$9 < \Delta L_{Ref} \leq 12$	5	$8 < \Delta L_{Ref} + \delta L \leq 10$	5
$12 < \Delta L_{Ref}$	6	$10 < \Delta L_{Ref} + \delta L$	6

6.3. Methode

Die im Folgenden präsentierten Messungen dienen zum einen dazu, das vorgeschlagene Verfahren auszuprobieren und zu validieren. Zum anderen werden explorativ potentielle Parameter mit Einfluss auf die THK näher untersucht. Auf diese Weise werden unterschiedliche Ziele verfolgt und gemeinsam bearbeitet.

6.3.1. Probanden

An den folgenden Experimenten nehmen insgesamt 14 Probanden im Alter zwischen 22 und 30 Jahren (Median 26,5 Jahre) freiwillig teil. Drei Personen sind weiblich, elf männlich. Es handelt sich um Mitglieder der Arbeitsgruppe Akustik, Fachbereich Physik und dem „Graduiertenkolleg Psychoakustik“ an der Universität Oldenburg.

Die Probanden haben einschlägige Erfahrungen mit psychoakustischen Messungen und Messverfahren. Sie sind normalhörend in dem Sinne, dass sie ihr eigenes Hörvermögen als „normal“ einschätzen und keine Hörstörungen bekannt waren oder sind. Bei Zweifel oder Unsicherheit wird zur Klärung ein Audiogramm aufgenommen. Weitere Tests der allgemeinen Hörfähigkeit werden nicht durchgeführt. Die Probanden werden für ihre Mitarbeit nicht entlohnt.

6.3.2. Anwendung und Konfiguration des Verfahrens

Für die folgenden Untersuchungen wird das oben vorgestellte 2I-2AFC 1-up 1-down Verfahren benutzt. Als Vergleichsstimulus wird das THK Referenzsignal (700 Hz Sinuston in GAR) verwendet. Die komplette Versuchssteuerung erfolgt

mit der Software SI [PK96], die auf einer Silicon Graphics (SGI) Workstation implementiert ist.

Die Generierung der Stimuli erfolgt während des Experiments durch SI mit 16 Bit Amplitudendynamik und einer Abtastfrequenz f_S von 22 050 Hz. Die Stimuli werden den Probanden⁴⁵ über einen internen D/A-Wandler (Ariel-Karte) diotisch⁴⁶ per Kopfhörer (Sennheiser HDA200) in einer schallisolierten Kabine präsentiert.

Der Gesamtpegel des GAR beträgt konstant $L_{GAR}=50$ dB SPL. Der Pegel des Tones im THK Referenzsignal wird hingegen während der Messung adaptiv verändert, so dass die bloße Angabe des Tonpegels $L_{T_{Ref}}$ keinen Sinn ergibt, vgl. Abbildung 6.1. Stattdessen wird der *Startpegel* des Tones angegeben.

Die Messungen werden jeweils für zwei verschiedene Startpegel durchgeführt. Diese werden als ± 10 dB relativ zu 45 dB SPL gewählt. Die beiden Startpegel betragen somit 55 dB SPL bzw. 35 dB SPL.

Für den kleineren Startpegel von 35 dB SPL liegt die maßgebliche Differenz ΔL_{Ref} etwa bei 1 dB, d.h. der Ton im Referenzsignal ist nahezu vollständig maskiert und nur schwach wahrnehmbar. Für den größeren Startpegel von 55 dB SPL ist die maßgebliche Differenz ca. 21 dB und der Ton damit sehr prominent.

Diese beiden Startpegel sollen gewährleisten, dass sich die Beurteiler einmal der gesuchten THK zunächst adaptiv „von unten“, im anderen Fall „von oben“ annähern. Die beiden Durchgänge werden in einer über die Probanden ausbalancierten Reihenfolge gemessen, um so den Einfluss des Startpegels zu minimieren.

Die verschiedenen Ausprägungen der untersuchten Parameter der THK werden gleichzeitig mit *interleaved Tracks* gemessen, d.h. der Vorgang der adaptiven Beurteilung findet simultan für mehrere Testsignale (*Tracks*) gleichzeitig statt.

Da dem Beurteiler (zu) viele Tracks intermittierend dargeboten werden, gelingt es ihm nicht, dem Verlauf der adaptiven Beurteilung für die einzelnen Tracks zu folgen. Der Proband kann maximal das aktuelle Trial anhand des zuvor bewerteten Trials beurteilen, er ist aber in der Regel nicht in der Lage, die einzelnen Tracks für die verschiedenen Ausprägungen des untersuchten Parameters nachzuverfolgen.

Darüber hinaus bekommen die Probanden keinerlei Feedback bzgl. ihrer Urteile/Bewertungen. Auf diese Weise werden kognitive Einflüsse minimiert und die Probanden gezwungen, sich bei jeder neuen Bewertung ausschließlich auf ihre Perzeption zu konzentrieren. Beide Maßnahmen dienen in erster Linie dazu, systematische Abweichungen bei der Bewertung der Stimuli zu minimieren.

Die Anfangsschrittweite für den Pegel des Tons im Vergleichsstimulus (d.h. im Referenzsignal) ist zunächst 8 dB. Die Schrittweite wird dann halbiert auf

⁴⁵ Im Sinne der Norm werden die Probanden synonym als *Beurteiler* bezeichnet.

⁴⁶ Eine diotische Darbietung liegt vor, wenn ein identisches Signal auf beiden Ohren synchron präsentiert wird.

4 dB und schließlich ist die minimale Schrittweite $S_{Min}=2$ dB. Während der Messphase werden sechs weitere Umkehrpunkte ermittelt, aus denen der Median samt Interquartilbereich als Messergebnis berechnet wird.

Die Genauigkeit wird bei dieser Konfiguration zugunsten schnellerer Durchführbarkeit bewusst gering angesetzt. Auf diese Weise kann eine größere Anzahl verschiedener Messungen (Parameter) mit vertretbarem zeitlichen Aufwand für die Probanden durchgeführt werden.

Die Länge der Test- und Vergleichsstimuli beträgt jeweils 1000 ms, die Pause zwischen den beiden Signalen 500 ms. Nachdem der Proband sein Urteil abgegeben hat, erfolgt automatisiert die nächste Signaldarbietung ohne zusätzliche Pausen.

Alle Probanden kennen die Definition der THK entsprechend der DIN 45681. Zusätzlich wird jeder Teilnehmer mindestens einmal bzgl. der Orientierung zur THK (Kapitel 4.7) aufgeklärt.

Die Aufgabe der Beurteiler bei allen Messungen besteht darin, zu beurteilen, welcher der paarweise dargebotenen Stimuli tonhaltiger ist⁴⁷.

Im Vorfeld dieser Untersuchung wird die voranstehende Konfiguration mittels einer unveröffentlichten Pilotstudie [Rie00] evaluiert. Die damit gewonnenen Daten zeigen, dass das hier vorgestellte Verfahren samt der beschriebenen Konfiguration geeignet ist, Parameter der THK zu untersuchen.

6.3.3. Darstellung der Ergebnisse

Das Ergebnis jeder einzelnen adaptiven Messung ist der Tonpegel des THK Referenzstimulus am PSE L_{TPSE} , d.h. derjenige Pegel, für den das Referenzsignal (Vergleichsstimulus) im Mittel die gleiche THK hervorruft wie das zu testende Signal (Teststimulus).

Der Tonpegel L_{TPSE} wird von der Software zur Versuchssteuerung am Ende einer Messung als Median über die Umkehrpunkte der Messphase berechnet. Das Skalenniveau der so durchgeführten Messung liegt somit oberhalb vom ordinalen Niveau, de facto wird aber keine Intervallskala erreicht. Aus diesem Grund erfolgen alle weiteren Auswertungen mit parameterfreien Methoden, solange keine anders lautenden Angaben gemacht werden.

Die folgenden Messergebnisse werden auf die gleiche Art und Weise dargestellt. Ein Beispiel dazu zeigt Abbildung 6.5. Auf der Abszisse sind jeweils als unabhängige Variablen die gemessenen Ausprägungen des Teststimulus aufgetragen. Auf den beiden Ordinaten ist ein Maß für die THK dargestellt, das jeweils unterschiedlich skaliert vom Tonpegel am PSE L_{TPSE} abhängt.

Auf der rechten Ordinate ist die maßgebliche Differenz ΔL_{RefPSE} für den Vergleichsstimulus (d.h. das Referenzsignal) am PSE nach DIN 45681:2006-08

⁴⁷ Die konkrete Anweisung bzw. Frage lautet: „Welches Signal ist tonhaltiger?“

aufgetragen. Sie wird berechnet, indem der Frequenzgruppenpegel um den Referenzton für das GAR nach Gleichung 6.1 zur Berechnung der maßgeblichen Differenz (Gleichung 6.4) eingesetzt wird:

$$(6.8) \quad \Delta L_{RefPSE} = L_{RefPSE} - (L_{GAR} - 10 \log(N_{FG})\text{dB}) - a_v(f_{T_{Ref}}).$$

Falls gewünscht, kann der Tonzuschlag für die so ermittelten maßgeblichen Differenzen ΔL_{RefPSE} anhand von Tabelle 6.1 direkt abgelesen werden.

Für die linke Ordinate wird die Pegeldifferenz L_D zwischen der maßgeblichen Differenz des Vergleichsstimulus am PSE ΔL_{RefPSE} und der nominellen maßgeblichen Differenz des Teststimulus ΔL_{Test} definiert:

$$(6.9) \quad L_D := \Delta L_{RefPSE} - \Delta L_{Test}.$$

Wenn für beide Stimuli die Definition der maßgeblichen Differenz in obige Formel eingesetzt wird, heben sich die entsprechenden Terme für das Rauschen gegenseitig auf:

$$(6.10) \quad L_D = \left(L_{RefPSE} - a_v(f_{T_{Ref}}) \right) - \left(L_{Test} - a_v(f_{T_{Test}}) \right).$$

Bis auf eine Ausnahme entspricht in den folgenden Messungen die Frequenz (bzw. die hervorgerufene Tonhöhe) des Teststimulus $f_{T_{Test}}$ der Frequenz (Tonhöhe) des Referenzstimulus $f_{T_{Ref}}$. Somit ist das Verdeckungsmaß $a_v(f)$ für beide identisch und fällt aus Gleichung 6.10 heraus:

$$(6.11) \quad L_D = L_{RefPSE} - L_{Test}.$$

Gleichung 6.11 ist leicht zu interpretieren: Eine positive Differenz $L_D > 0$ gibt an, dass die beurteilte THK des Referenzstimulus am PSE (d.h die THK für den Teststimulus) größer ist als die nominelle THK des Teststimulus. Für $L_D < 0$ ist die beurteilte THK kleiner als die nominelle THK.

Die beiden Ordinaten unterscheiden sich um einen konstanten Pegelwert. Um diesen zu ermitteln, wird Gleichung 6.9 nach ΔL_{RefPSE} freigestellt und die Definition für die maßgebliche Differenz ΔL_{Test} analog zu Gleichung 6.8 ersetzt:

$$(6.12) \quad \Delta L_{Test} = - (L_{GAR} - 10 \log(N_{FG})\text{dB}) - a_v(f_{T_{Ref}}) - L_D.$$

Für den hier verwendeten Frequenzbereich des GAR von 0 Hz bis zur halben Abtastrate von 11 025 Hz ist nach Zwicker und Fastl [ZF90] (Tabelle 6.1) die Anzahl der Frequenzgruppen $N_{FG} \approx 22,7$. Der Gesamtpegel des GAR beträgt $L_{GAR}=50$ dB, der mittlere Startpegel $\bar{L}_{T_{start}}=45$ dB. Das Verdeckungsmaß wird mit Gleichung A.1 zu $a_v(700\text{ Hz})=-2,5$ dB abgeschätzt. Damit ergibt sich die gesuchte Differenz zu:

$$(6.13) \quad \Delta L_{RefPSE} = L_D + (45 - 50 + 13,6 + 2,5)\text{dB} = L_D + 11,1\text{ dB}.$$

Die dünnen (farbigen) Linien in den folgenden Abbildungen der durchgeführten Messungen repräsentieren die individuellen Daten der beiden Messreihen für die verschiedenen Startpegel. Dabei wird pro Proband jeweils eine andere Farbe benutzt, die Messungen für den ersten Startpegel sind mit einem Stern (*) gekennzeichnet, die Messungen für den zweiten Startpegel durch eine Box (\square). Die einzelnen Datenpunkte sind für eine bessere Darstellung entlang der Abszisse leicht verschoben.

Die Messungen für die beiden Startpegel werden für jeden Proband zunächst individuell als Median zusammen gefasst⁴⁸. Mit den so ermittelten Medianen werden die statistischen Prüfungen der Daten vorgenommen.

Im Anhang F.3 werden die Daten zusätzlich auf Intervallskalenniveau untersucht. Dort wird zunächst statistisch überprüft, ob ein Effekt des Startpegels vorliegt und ob eine Zusammenfassung der Daten für die verschiedenen Startpegel gerechtfertigt ist.

Zur interindividuellen Zusammenfassung der Daten wird der Median samt Interquartilbereich verwendet. Dieser ist in den Abbildungen als dickere, schwarze Linie gekennzeichnet.

Einige der extrem ausgeprägten individuellen Probandendaten liegen außerhalb des dargestellten Bereiches. Dieser Nachteil wird zugunsten einer einheitlichen Darstellung aller folgenden Abbildungen als vertretbar angesehen.

6.4. Messungen

6.4.1. THK als Funktion der Anzahl von Tönen

Das Ziel dieser Untersuchung besteht darin, zu untersuchen, wie sich die THK vergleichbar großer tonaler Komponenten aggregiert.

Dazu werden harmonische komplexe Töne⁴⁹ (HKT) mit variierender Anzahl an Teiltönen k in einem rauschhaften Hintergrund verwendet. Die HKTs werden während des Experiments⁵⁰ generiert. Die Grundfrequenz f_G der HKTs beträgt für alle in diesem Kapitel präsentierten Messungen 700 Hz, d.h. sie entspricht der Frequenz des Referenztones.

Es werden vier verschiedene HKTs mit $k=1, 2, 4, 8$ Einzeltönen untersucht. Der HKT mit $k=1$ ist ein einzelner (reiner) Sinuston, die HKTs mit $k>1$ enthalten jeweils alle Harmonischen bis zum Teilton k , vgl. Abbildung 6.2.

Der Abstand zwischen den einzelnen Teiltönen ist für die niedrigeren Harmonischen deutlich größer als eine Frequenzgruppenbreite, geht jedoch für die

⁴⁸ Da nur zwei Messwerte (Startpegel) pro Proband vorliegen, reduziert sich die Bildung des Medians in diesem Fall auf eine Mittelwertbildung.

⁴⁹ Bezeichnung nach Terhardt [Ter98].

⁵⁰ Für die Generierung wird der interne Befehl HTC der Software SI [PK96] benutzt.

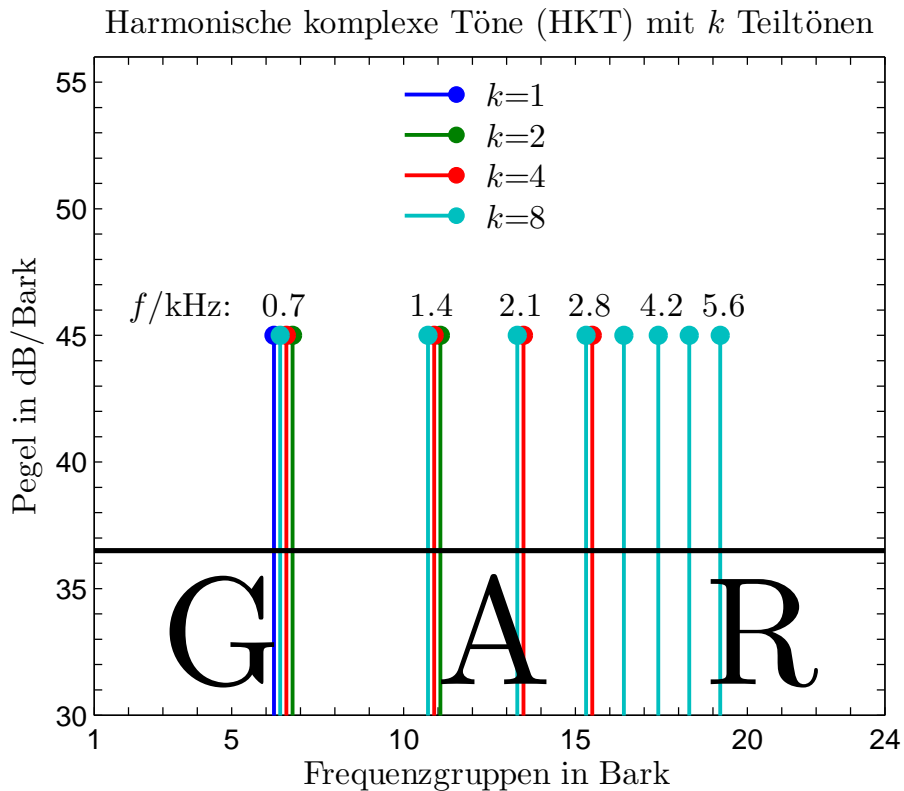


Abbildung 6.2.: Schematische Darstellung der HKTs als Dichtespektrum. Auf der Abszisse sind Frequenzgruppen in Bark aufgetragen, die Ordinate gibt den Pegel der Teiltöne pro Frequenzgruppe in dB/Bark an.

In der Mitte der Abbildung ist zusätzlich die Frequenz für die Teiltöne in kHz aufgelistet.

Die horizontale Linie bei 36,4 dB symbolisiert den Frequenzgruppenpegel des gleichmäßig anregenden Rauschens (GAR).

höheren Harmonischen ungefähr auf eine Frequenzgruppenbreite zurück.

Der Pegel $L_{HTK(k=1)}$ des Grundtons sowie aller einzelnen Teiltöne des HKT beträgt 45 dB SPL. Der Gesamtpegel für die Teststimuli $L_{HTK(k)}$ ergibt sich daher zu

$$(6.14) \quad L_{HTK(k)} = L_{HTK(k=1)} + 3 \log_2(k) \text{ dB.}$$

Bei Verdopplung der Anzahl der Teiltöne k erhöht sich der Gesamtpegel des HKT demnach jeweils um 3 dB⁵¹.

⁵¹ Dieser Zusammenhang wird hier als *ohne „Energieausgleich“* bezeichnet, da im Weiteren auch HKTs *mit „Energieausgleich“* untersucht werden, vgl. Abschnitt 7.1 und insbesondere Abbildung 7.1.

6. Einfluss verschiedener Signalparameter auf die Tonhaltigkeit

Als Hintergrundrauschen wird, so wie für das THK Referenzsignal, gleichmäßig anregendes Rauschen (GAR) verwendet. Auf diese Weise kann die THK der vier HKTs separat analysiert werden, ohne den Einfluss des Rauschens besonders berücksichtigen zu müssen.

Der Gesamtpegel für das Rauschen L_{GAR} beträgt 50 dB SPL. Der Frequenzgruppenpegel L_G für das GAR L_{FGGAR} ergibt sich demnach mit Gleichung 6.1 zu 36,4 dB, siehe Abbildung 6.2. Weitere Einzelheiten sind Anhang F.2 zu entnehmen.

Ergebnisse

Die Ergebnisse sind in der im Abschnitt 6.3.3 angekündigten Weise in Abbildung 6.3 dargestellt.

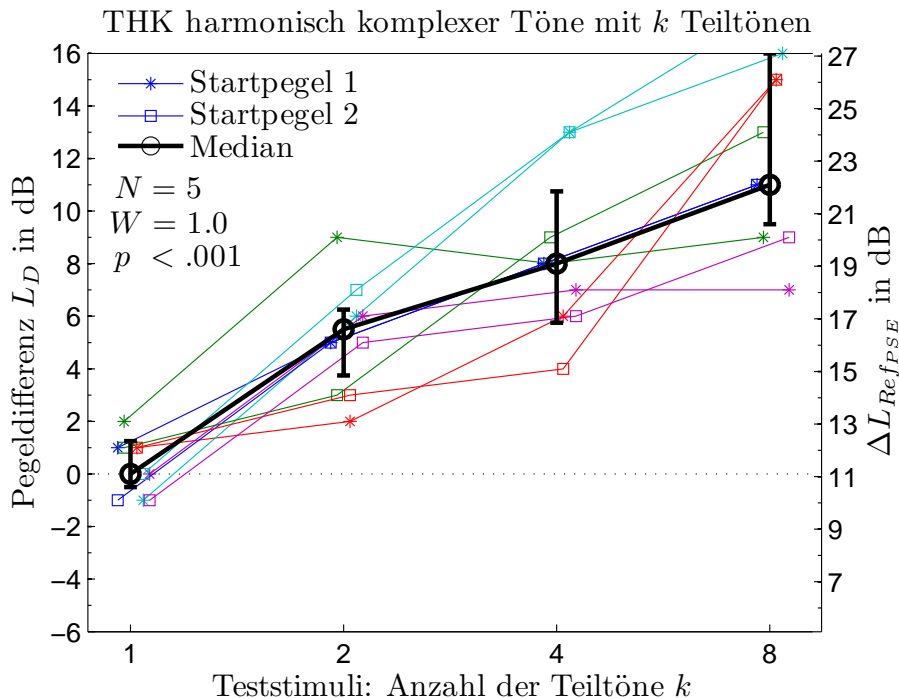


Abbildung 6.3.: THK harmonisch komplexer Töne (HKT) ohne Energieausgleich für fünf Probanden. Aufgetragen ist die THK (L_D , bzw. ΔL_{RefPSE} siehe Abschnitt 6.3.3) über die Anzahl der Teiltöne k des Teststimulus.

Zur parameterfreien Überprüfung der Konkordanz, d.h. der Übereinstimmung zwischen den $N=5$ Beurteilern, wird *Kendalls Konkordanzkoeffizient* W herangezogen. Dieser Konkordanzkoeffizient kann Werte zwischen $W=1$ (perfekte Übereinstimmung) und $W=0$ (keine Übereinstimmung) annehmen. Für die vor-

liegenden Daten zeigt sich eine perfekte Übereinstimmung der Urteiler, d.h. $W=1.0$. Dieses Ergebnis ist mit $p<.002$ signifikant.

Abbildung 6.3 zeigt, dass die wahrgenommene THK monoton mit steigender Anzahl von Einzelkomponenten k des HKT anwächst. Der Friedman-Test ($\chi^2(3, N=5)=13.0$) bestätigt diesen Befund mit einer exakt berechneten Signifikanz von $p<.002$.

Um die Unterschiede zwischen den einzelnen HKTs zu untersuchen, wird der Wilcoxon-Test für verbundene Stichproben paarweise für die entsprechenden Kombinationen ausgeführt. Die Ergebnisse dieser Tests sind in Tabelle 6.2 zusammengefasst.

Tabelle 6.2.: Paarweise Wilcoxon-Tests für die THK der HKT-Stimuli. Angegeben ist von links nach rechts das Stimulus-Paar, der Z-Wert, die unkorrigierte Signifikanz (exakt, zweiseitig) und die Signifikanz (exakt, zweiseitig) bei Korrektur des α -Fehlers (Bonferroni).

Paar	Z-Wert	Signifikanz, 2-seitig unkorrigiert	Signifikanz, 2-seitig korrigiert (Bonferroni)
$k=1 ; k=2$	-2,02	.063	.38
$k=1 ; k=4$	-2,02	.063	.38
$k=1 ; k=8$	-2,02	.063	.38
$k=2 ; k=4$	-2,02	.063	.38
$k=2 ; k=8$	-2,02	.063	.38
$k=4 ; k=8$	-2,02	.063	.38

Der Wilcoxon-Test zeigt tendenzielle, aber keine signifikanten Unterschiede zwischen allen möglichen Kombinationen von Stimuli. Bei Berücksichtigung der Inflation des α -Fehlers sind definitiv keine Unterschiede zwischen den Stimuli nachweisbar.

Die Diskussion der Ergebnisse des nächsten Abschnitts (6.4.1) vorwegnehmend, ist ein Anstieg der THK mit steigender Anzahl der Teiltöne zu erwarten. Dementsprechend kann eine gerichtete Hypothese aufgestellt werden, die eine einseitige statistische Prüfung dieses Sachverhalts erlaubt. Die Signifikanz der dritten Spalte aus Tabelle 6.2 halbieren sich damit zu $p=.031$ und sind signifikant.

Einer Faustregel von Roberts und Russo [RR99] entsprechend, braucht die Inflation des α -Fehlers für sukzessive Tests nicht berücksichtigt werden, wenn ein Vergleich weniger durchgeführt wird als Faktorstufen existieren. Es ist daher in diesem Sinne legitim zu behaupten, dass sich die drei Paare folgender HKTs signifikant unterscheiden: HKTs mit $k=1$ und $k=2$ ($p=.031$), $k=2$ und $k=4$ ($p=.031$) sowie $k=4$ und $k=8$ ($p=.031$).

Über die Rechtfertigung dieser post-hoc Abschätzung kann, besonders vor dem Hintergrund des geringen Stichprobenumfangs, gegensätzlich argumentiert werden. Die im weiteren gefundenen Ergebnisse (vgl. Kapitel 7) zeigen jedoch die Plausibilität der Interpretation der vorliegenden Daten.

Unter der Annahme von Intervallskalenniveau dieser Daten⁵² kann eine ANOVA durchgeführt werden. Die Ergebnisse dafür sind im Anhang F.3.1 ausführlich dargestellt.

Die Ergebnisse der ANOVA bestätigen alle zuvor gefundenen Aussagen. Insbesondere deckt die ANOVA eine sehr hohe Effektgröße von $\eta^2=.863$ auf, die auf eine große Teststärke hindeutet. Damit reichen sehr wenige Probanden aus, um ein signifikantes Ergebnis zu erhalten. Das zeigt sich auch an den zuvor präsentierten Daten für nur fünf Probanden.

Diskussion

In der DIN 45681E:1992-01 wird angenommen, dass nur die tonale Komponente⁵³ mit der größten maßgeblichen Differenz für die THK entscheidend ist. Für die verwendeten Stimuli hat jeder Teilton analog zu Gleichung 6.6 eine identische maßgebliche Differenz ΔL_{92} von 14,6 dB. Aus diesem Grund sollte die THK für alle Stimuli im Wesentlichen gleich sein.

In der neuesten Version der Norm DIN 45681:2006-08 hängt die maßgebliche Differenz ΔL jedoch über das Verdeckungsmaß $a_v(f)$ von der Frequenz jedes einzelnen Teiltönen ab. Sie liegt zwischen 11,1 dB für 700 Hz und 13,2 dB für 5600 Hz⁵⁴. Unter der Voraussetzung, dass die acht Teiltöne der HKTs in verschiedenen Frequenzgruppen liegen, sollte die THK daher für die vier Stimuli mit wachsender Anzahl der Teiltöne ebenfalls ansteigen. Der Unterschied zwischen der maßgeblichen Differenz für den kleinsten Teilton zum größten Teilton beträgt lediglich 2,1 dB, der Anstieg der wahrgenommenen THK vom Stimulus mit $k=1$ zum Stimulus mit $k=8$ sollte nach DIN 45681:2006-08 etwa im Bereich dieses Wertes liegen.

Tatsächlich ist der Anstieg der THK bei anwachsender Anzahl von Teiltönen bedeutend größer. Er beträgt hier etwa 5 dB für die erste Verdopplung der Teiltonanzahl ($k=1 \rightarrow k=2$). Damit ist der Anstieg größer als aufgrund einer Pegeladdition der beiden Teiltöne zu erwarten wäre (eine Intensitätsverdopplung entspricht 3 dB).

Für eine subjektive Lautheitsverdoppelung wäre ein Pegelanstieg von ca. 10 dB notwendig [FZ07]. Der vorliegende Effekt liegt gewissermaßen zwischen

⁵² Das Skalenniveau der Daten wird im Anhang F.3 diskutiert.

⁵³ Die acht Teiltöne des HKT(8) entsprechen acht tonalen Komponenten bei Anwendung der DIN 45681.

⁵⁴ Die maßgeblichen Differenzen für die Teiltöne sind Tabelle F.2 im Anhang F.2 zu entnehmen.

diesen beiden Abhängigkeiten. Hansen [Han10] untersucht diesen Sachverhalt, indem er die THK der Teiltöne ihren Teillautheiten gegenüberstellt. Seine Hypothese lautet, dass die Aggregation der THK auf die Summation der Teillautheiten der Teiltöne zurück geführt werden kann.

Diese Hypothese und allgemein die Fragestellung der Aggregation der THK von Tönen bzw. Teiltönen wird im Kapitel 7 systematisch untersucht. Die gefundenen Zusammenhänge werden dort anhand weiterer Messungen vertieft und eingehend diskutiert, so dass an dieser Stelle keine weitere Interpretation der Daten erfolgt.

Es bleibt abschließend festzuhalten, dass die Konkordanz mit $W=1$ für die Teilnehmer dieser Untersuchung bemerkenswert hoch ist und selbst für die geringe Anzahl von fünf Beurteilern die Unterschiede zwischen den vier HKTs nachgewiesen werden können. Das ist vorrangig auf die beachtliche Effektgröße von $\eta^2=.863$ zurückzuführen, die es hier ermöglicht, mit einer geringen Anzahl von Probanden aussagekräftige Daten zu erhalten.

6.4.2. THK als Funktion der Harmonizität von Teiltönen

Die Stimuli im vorangegangenen Abschnitt sind allesamt „harmonisch“ in dem Sinne, dass der Grundton der HKTs zugleich der kleinste im Signal enthaltene Teilton ist (vgl. Terhardt Kapitel 3.5.1 [Ter98]). Um zu untersuchen, ob der gefundene Effekt der Anzahl der Teiltöne auch für nicht harmonische Tonkomplexe gültig ist, wird der HKT mit $k=2$ aus dem vorangegangenen Experiment auf der Frequenzachse verschoben.

Es werden vier Verschiebungen des HKT($k=2$) vorgenommen: $f_{shift}=-300$ Hz, -150 Hz, 150 Hz und 300 Hz. Die so erzeugten Stimuli zeigt Abbildung 6.4. Der unverschobene HKT($k=2$) ist rot eingezeichnet. Die vier Verschiebungen sind durch Pfeile angedeutet, die daraus entstehenden VHKTs werden jeweils durch eine andere Farbe repräsentiert.

Die Verschiebung bewirkt, dass der Grundton nicht mehr zugleich der kleinste im Signal enthaltene Teilton ist. Für die fünf verwendeten Stimuli ergibt sich der Grundton als der größte gemeinsame Teiler der Teiltöne zu 100 Hz, 50 Hz, 700 Hz, 50 Hz und 100 Hz. Diese Frequenzen entsprechen im weitesten Sinne virtuellen Tonhöhen der Stimuli [Vor95].

Bis auf die Verschiebung der Teiltöne werden die Stimuli genauso wie in der vorhergehenden Untersuchung generiert. Das GAR sowie sämtliche Pegel bleiben unverändert.

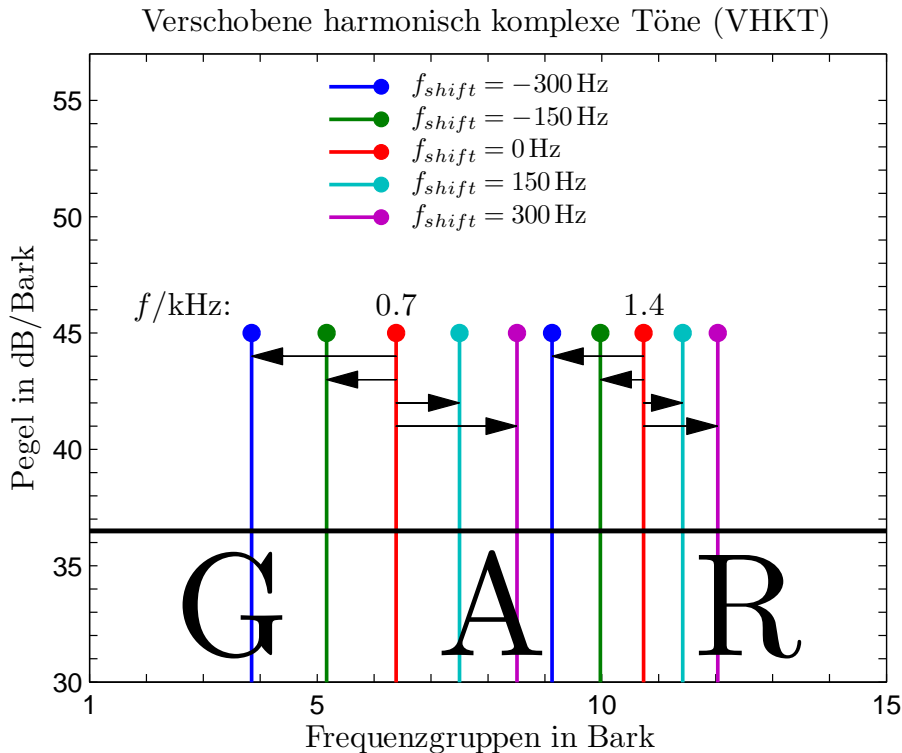


Abbildung 6.4.: Schematische Darstellung der verschobenen harmonischen komplexen Töne (VHKT) als Dichtespektrum. Auf der Abszisse sind Frequenzgruppen in Bark aufgetragen, in der Mitte der Abbildung zusätzlich die Frequenzen der Teiltöne für den nicht verschobenen Stimulus in kHz.

Die horizontale Linie bei 36,4 dB symbolisiert den Pegel des gleichmäßig anregenden Rauschens (GAR).

Ergebnisse

Um die Vergleichbarkeit dieser Untersuchung mit der zuvor beschriebenen zu gewährleisten, nehmen die gleichen fünf Probanden an den Bewertungen teil.

Die gemessenen Daten sind wie zuvor in Abbildung 6.5 dargestellt. Die weitere Analyse erfolgt wie in Abschnitt 6.3.3 erläutert.

Der Friedman-Test ($\chi^2(4, N=5)=6.3$) zeigt keinen signifikanten Einfluss des Parameters Frequenzverschiebung auf die wahrgenommene THK ($p=.17$). Wie der Abbildung zu entnehmen ist, streuen die individuellen Daten mehr oder weniger zufällig um den Median. Die Übereinstimmung der Urteiler ist mit $W=.32$ eher schlecht und nicht signifikant auf dem 5%-Niveau.

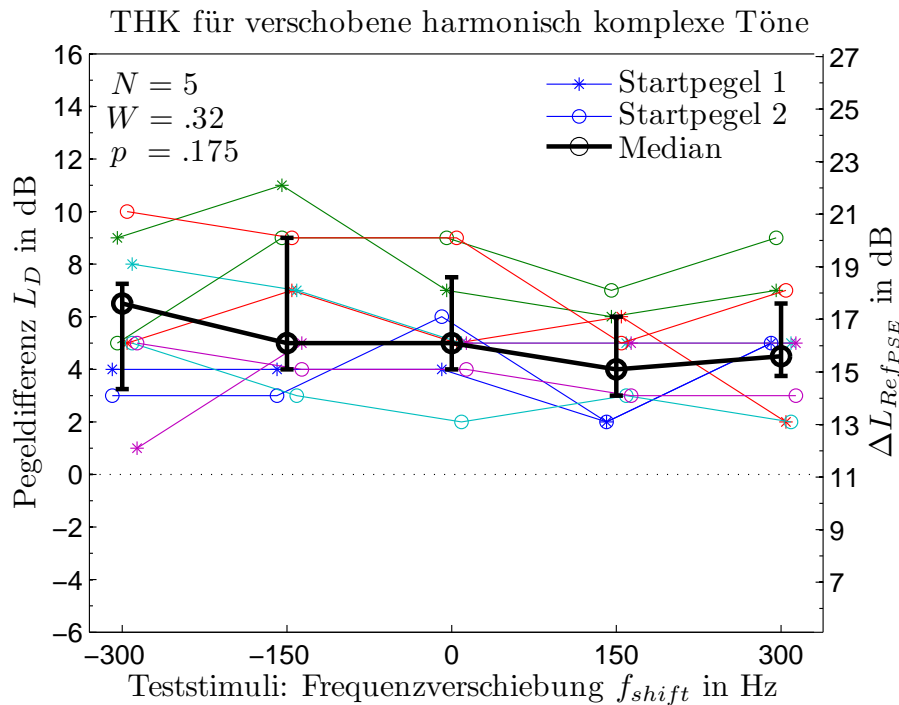


Abbildung 6.5.: Aufgetragen ist die Pegeldifferenz L_D und die maßgebliche Differenz ΔL_{RefPSE} (vgl. Abschnitt 6.3.3) der fünf Teststimuli über die Frequenzverschiebung der Teiltöne f_{shift} für fünf Probanden.

Diskussion

Es zeigt sich keine Abhängigkeit der THK von der Harmonizität der Teiltöne. Die Harmonizität ist somit nicht der entscheidende Faktor für die zuvor gefundenen Effekte der HKTs auf die THK.

Ein Vergleich der THK der VHKTs mit dem HKT($k=2$) ergibt eine gute Übereinstimmung: Der Median der Pegeldifferenz L_D für die VHKTs ist 5 dB mit einem Interquartilbereich zwischen 4 dB und 7 dB. Demgegenüber beträgt die THK für den HKT($k=2$) 5,5 dB bei einem Interquartilbereich zwischen 3,75 dB und 6,25 dB.

Unter der Annahme intervallskalierter Daten wird im Anhang F.3.2 ergänzend zu den diskutierten Analysen eine zweifaktorielle ANOVA mit Messwiederholungen durchgeführt. Die ANOVA bestätigt die Ergebnisse der parameterfreien Analyse. Der Mittelwertvergleich ergibt für die VHKTs eine THK von $L_D=5,3$ dB bei einer Standardabweichung von 1,8 dB gegenüber einer THK von $L_D=(5,1 \pm 1,6)$ dB für den HKT($k=2$).

Die gefundenen Resultate sind in Einklang mit der DIN 45681:2006-08, so lange die Teiltöne nicht in die gleichen Frequenzgruppen fallen. Um diese Beschrän-

kung bzgl. mehrerer Teiltöne in der gleichen Frequenzgruppe zu untersuchen, wird im nächsten Abschnitt der Abstand zwischen den Teiltönen variiert und bewertet.

6.4.3. THK als Funktion des Abstandes zweier Sinustöne

Nachfolgend wird der Einfluss des spektralen Abstandes zweier Sinustöne auf die THK untersucht. Die Generierung der Stimuli erfolgt, indem zwei Sinustöne mit einem Frequenzabstand von 2 Hz, 10 Hz, 70 Hz, 140 Hz und 270 Hz symmetrisch um die Frequenz 700 Hz angeordnet werden. Die Pegel und das Rauschen sind so wie in den vorangegangenen Untersuchungen verwendet.

Ergebnisse

Abbildung 6.6 zeigt die gemessenen Daten für die THK.

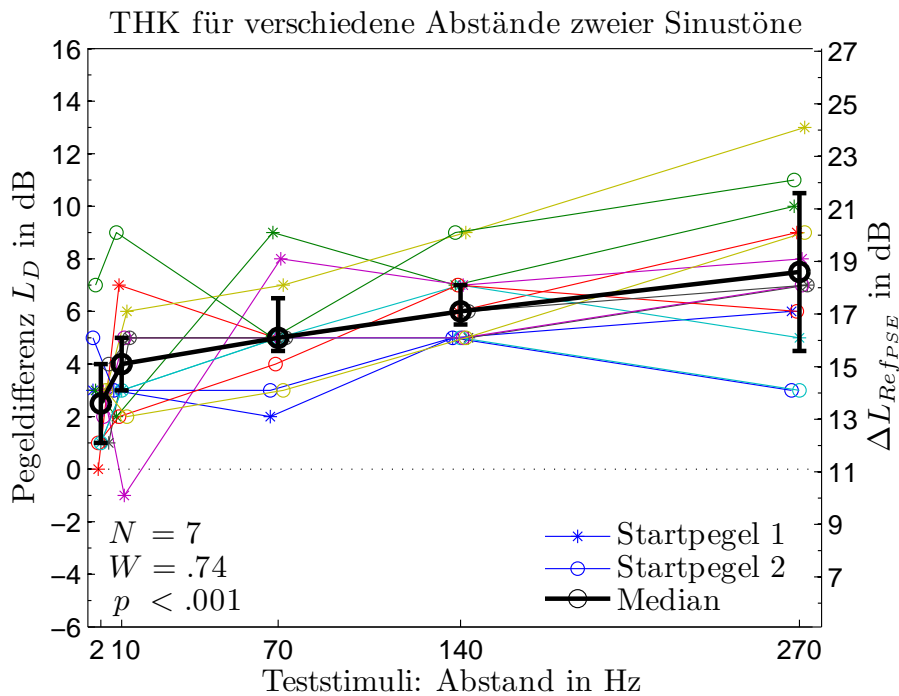


Abbildung 6.6.: Einfluss des Abstandes zweier Sinustöne auf die THK für sieben Probanden. Die Art der Darstellung ist analog zu den vorangegangenen Abbildungen, vgl. auch Abschnitt 6.3.3.

Die Übereinstimmung der Probanden ist mit $W=.0737$ ($p<.001$) akzeptabel, der Friedman-Test ($\chi^2(4, N=7)=20.4, p<.001$) zeigt eine deutliche Abhängigkeit der THK vom Abstand der beiden Sinustöne.

Der Wilcoxon-Test (exakt, zweiseitig, unkorrigiert) zeigt signifikante Unterschiede des Abstands von 2 Hz zu 70 Hz ($p=.031$), 140 Hz ($p=.031$) und 270 Hz ($p=.016$). 10 Hz unterscheidet sich signifikant von 140 Hz ($p=.016$) und 270 Hz ($p=.016$). Der Abstand 70 Hz unterscheidet sich signifikant von 140 Hz ($p=.047$) und 270 Hz ($p=.047$). Bei Berücksichtigung der Inflation des α -Fehlers ist keiner der aufgeführten Unterschiede auf dem 5%-Niveau signifikant.

Unter der Annahme, dass die Ergebnisse des adaptiven THK-Messverfahrens Intervallskalenniveau besitzen, wird eine ANOVA durchgeführt, siehe Anhang **F.3.3**. Auch sie zeigt einen signifikanten Effekt des Abstandes ($F(4,24)=14.6$, $p<.001$). Post-hoc Tests (Bonferroni) ergeben signifikante Unterschiede auf dem 5%-Niveau für die folgenden Stimulus-Paare: 2 Hz und 140 Hz ($p=.009$), 2 Hz und 270 Hz ($p=.023$) so wie für 10 Hz und 270 Hz ($p=.011$).

Diskussion

Die DIN 45681:2006-08 gibt für die vorliegende Konfiguration der Teststimuli vor, dass die Tonpegel der beiden Teiltöne für Abstände unter ≈ 50 Hz (eigtl. 49 Hz, siehe Gleichung 2.4) energetisch addiert werden müssen, da sie nicht als zwei einzelne Töne aufgelöst werden können. Darüber werden die Töne getrennt wahrgenommen und es gilt die allgemeingültige Regel, dass nur der Teilton mit der größeren maßgeblichen Differenz weiter betrachtet wird, vgl. Kapitel 2.9.

Diese Annahme der Norm wird durch die gemessenen Daten nur teilweise bestätigt. Es zeigt sich, dass für sehr kleine Abstände der beiden Teiltöne die THK etwa einen Wert annimmt, wie er bei Addition der Teiltonpegel (+3 dB) zu erwarten wäre. Allerdings wird die THK für größere Abstände der Teiltöne nicht kleiner, sondern nimmt weiter zu.

Diese Tatsache ist gleichwohl nicht auf ein Problem der Norm für zwei eng benachbarte Teiltöne zurückzuführen, sondern auf den gleichen Widerspruch, der bereits in den beiden vorausgegangenen Untersuchungen festgestellt wurde. Bei der Aggregation von zwei Teiltönen mit nahezu gleicher maßgeblicher Differenz wird die THK unterschätzt.

Die auftretenden Schwebungen für kleine Abstände der Teiltöne scheinen keinen Einfluss auf die THK zu haben. Diese Hypothese kann an dieser Stelle jedoch nicht bewiesen werden.

Wird der Abstand der Teiltöne über eine Frequenzgruppenbreite hinaus ausgedehnt, ergeben sich ähnliche Bewertungen der THK wie in den vorangegangenen Untersuchungen. Für den Abstand von 140 Hz ist der Median der THK 6 dB (Interquartilbereich 5,5 dB bis 7 dB) gegenüber 5,5 dB (Interquartilbereich 3,75 dB bis 6,25 dB) für den HKT($k=2$). Diese Werte sind vergleichbar, zumal das Probandenkollektiv zwischen den beiden Untersuchungen nicht identisch ist.

6.4.4. THK als Funktion des absoluten Darbietungspegels

Im Folgenden wird die Abhängigkeit der THK vom absoluten Darbietungspegel für breitbandige Signale untersucht. Damit soll geklärt werden, ob es einen Unterschied macht, wenn derselbe Stimulus bei unterschiedlichen Pegeln dargeboten wird.

Die Hypothese lautet in diesem Fall, dass der absolute Pegel keinen Einfluss auf die THK hat, da sie im Wesentlichen nur von der maßgeblichen Differenz des Tones abhängt (DIN 45681:2006-08).

Als Basis-Stimulus wird hier der HKT($k=1$), d.h. der reine Ton aus der ersten Untersuchung, verwendet. Er wird mit den gleichen Pegeln ($L_{T_{rest}}=45$ dB SPL, $L_{GAR}=50$ dB SPL) wie bisher dargeboten.

Der Basis-Stimulus wird insgesamt um die Pegelwerte ± 10 dB, ± 20 dB und ± 30 dB abgesenkt bzw. angehoben. Auf diese Weise werden sieben Stimuli unterschiedlicher Intensität erzeugt. Für den schwächsten Stimulus beträgt das GAR lediglich 20 dB SPL und ist damit sehr leise, der stärkste Stimulus hat einen Rauschpegel von 80 dB SPL und ist vergleichsweise laut.

Ergebnisse

Die THK-Daten für acht Probanden sind in Abbildung 6.7 dargestellt. Auf Anhieb ist kein klarer Trend erkennbar, der Friedman-Test ($\chi^2(6, N=8)=23.12$) zeigt jedoch, dass der Darbietungspegel einen signifikanten Einfluss ($p < .001$) auf die THK besitzt.

Durch paarweise Anwendung des Wilcoxon-Test werden post-hoc signifikante Unterschiede (siehe Tabelle 6.3) zwischen einigen Stimuli gefunden. Bei Berücksichtigung der α -Fehler Inflation (Bonferroni) verlieren alle Unterschiede ihre statistische Aussagekraft.

Tabelle 6.3.: Relevante Wilcoxon-Tests für die sieben Darbietungspegel.

Paar	Z-Wert	Signifikanz (2-seitig)	Signifikanz, Bonferroni (2-seitig)
-30 ; -20	-2,1	.047	.99
-30 ; 0	-2,4	.016	.34
-30 ; 10	-2,5	.008	.16
-30 ; 20	-2,5	.008	.16
-20 ; 10	-2,1	.031	.65
-10 ; 0	-2,4	.016	.34
-10 ; 10	-2,2	.023	.48
-10 ; 20	-2,2	.031	.65

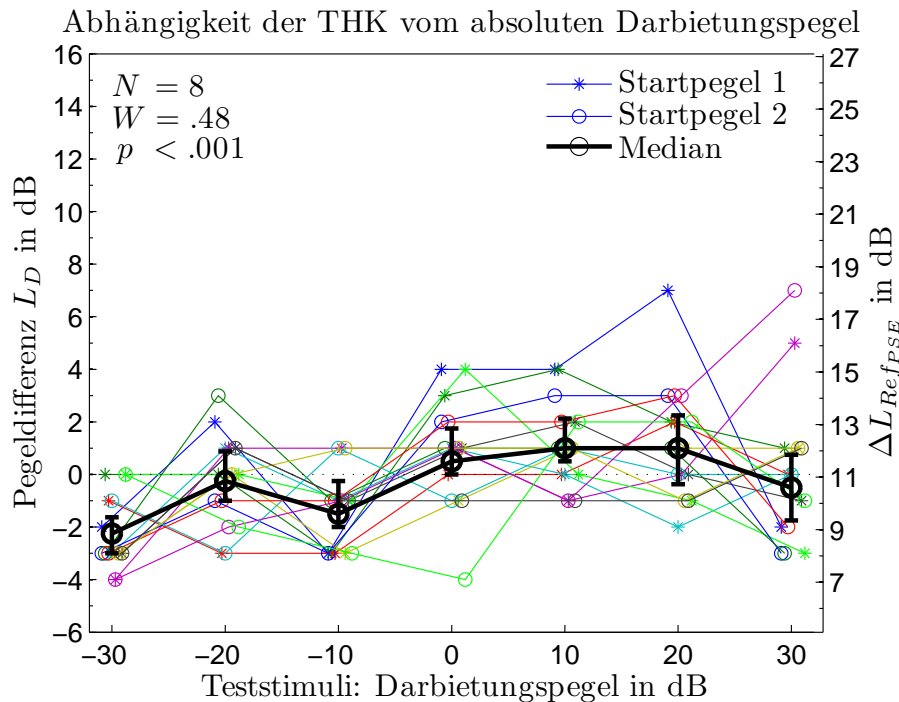


Abbildung 6.7.: Aufgetragen ist die THK (L_D , bzw. ΔL_{RefPSE} siehe Abschnitt 6.3.3) über den absoluten Darbietungspegel für acht Probanden.

Die Übereinstimmung der Probanden ist mit $W=.482$ mäßig. Sie wird im Wesentlichen durch den Stimulus mit dem Darbietungspegel von -30 dB hervorgerufen, der sich von fast allen anderen Stimuli unterscheidet.

Wird die Analyse ohne den Stimulus mit -30 dB wiederholt, so zeigt der Friedman-Test ($\chi^2(5, N=8)=13.94$) einen wesentlich schwächeren Effekt ($p=.016$) des Darbietungspegels, die Übereinstimmung der Probanden ($W=.349$) nimmt ab, d.h. die Probandenurteile werden weniger systematisch.

Unter der Annahme intervallskalierter Daten wird eine ANOVA durchgeführt (Anhang F.3.4). Sie offenbart eine signifikante Abhängigkeit der THK vom Darbietungspegel ($F(6,49)=5.5$, $p<.001$). Post-hoc Tests (Bonferroni) zeigen ähnliche Unterschiede zwischen den Stimuli wie die Wilcoxon-Tests, siehe Tabellen 6.3 und F.5.

Bei Anwendung der ANOVA auf die Daten ohne den Darbietungspegel -30 dB zeigt sich ein relativ schwacher Effekt des Darbietungspegels: ($F(5,42)=2.9$, $p=.025$). Post-hoc Vergleiche (Bonferroni) decken keine signifikanten Unterschiede zwischen einzelnen Darbietungspegeln auf.

Diskussion

Die aufgedeckten Effekte des Darbietungspegels auf die THK werden im Wesentlichen durch den leisesten Stimulus (Darbietungspegel -30 dB) hervorgerufen. Für diesen Stimulus hat das GAR nur einen Pegel von 20 dB SPL, d.h. nach Gleichung 6.1 ist der Frequenzgruppenpegel kleiner als 7 dB SPL.

Für derart kleine und damit schwelennahe Frequenzgruppenpegel kann nicht sichergestellt werden, dass die vorgenommene Kopfhörer-Entzerrung ausreichend ist. Es besteht die Möglichkeit, dass Teile der betreffenden Frequenzgruppe um 700 Hz unterhalb der Hörschwelle liegen. Das könnte z.B. bedeuten, dass der Pegel des Testtones dadurch verringert wird und die dem Probanden tatsächlich präsentierte maßgebliche Differenz für diesen Stimulus kleiner ist als angenommen.

Weiterhin wird die individuelle Hörschwelle der Probanden nicht ermittelt und berücksichtigt, sodass auch hier Effekte entstehen können, die nicht ausreichend kontrolliert werden.

Der Stimulus mit dem Darbietungspegel -30 dB muss daher kritisch betrachtet werden und sollte nicht vorbehaltlos in die Analyse mit einbezogen werden.

6.4.5. THK als Funktion der Bandbreite von Schmalbandrauschen

„Die Ausprägtheit eines Tones hängt von der Bandbreite des Tones und seiner Flankensteilheit ab“ (DIN 45681:2006-08 siehe auch Abschnitt 2.6).

Die Umsetzung dieser Erkenntnis in der Norm (Gleichung 2.2) wird überprüft, indem der $HKT(k=1)$ aus der ersten Untersuchung durch Schmalbandrauschen unterschiedlicher Bandbreite ersetzt wird. Die Pegel sind erneut so gewählt wie in den vorangegangenen Abschnitten (Schmalbandrauschen 45 dB SPL, GAR 50 dB SPL).

Die Flankensteilheit wird hier nicht näher untersucht, sondern als „unendlich steil“ angenommen. Das wird erreicht, indem das Schmalbandrauschen mittels einer inversen FFT aus einem vorgegebenen Spektrum erzeugt wird⁵⁵.

Die (lineare) Mittenfrequenz der Schmalbandrauschen liegt für alle Stimuli bei 700 Hz, die Bandbreiten betragen 2 Hz, 10 Hz, 70 Hz, 140 Hz und 270 Hz.

Die Hypothese ist, dass die THK für breiter werdende Schmalbandrauschen abnimmt.

⁵⁵ Die skizzierte Signalgenerierung wird durch den Befehl `BNOISE` der Software zur Versuchsteuerung (SI [PK96]) zur Verfügung gestellt.

Ergebnisse

Die gemessenen THK-Daten sind in Abbildung 6.7 dargestellt. Aus der Abbildung ist die Tendenz zu erkennen, dass die THK für ansteigende Bandbreiten des Schmalbandrauschens abfällt. Der Friedman-Test ($\chi^2(4, N=7)=19.26$) bestätigt diesen Effekt ($p < .001$). Die Übereinstimmung der Urteiler ist mit $W=.688$ verhältnismäßig hoch.

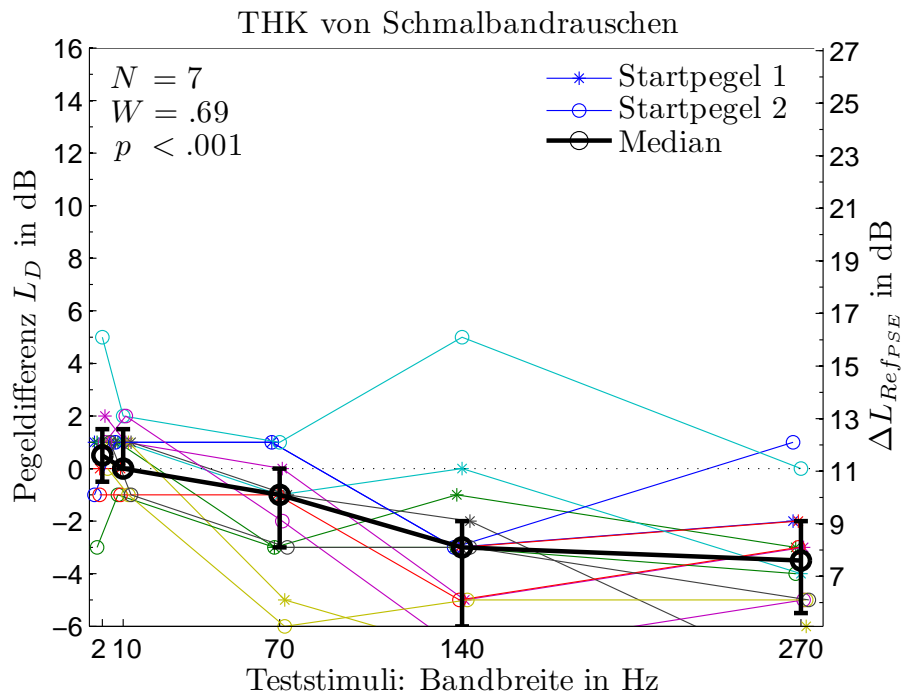


Abbildung 6.8.: Aufgetragen ist die THK (L_D , bzw. ΔL_{RefPSE} siehe Abschnitt 6.3.3) über die Bandbreite von Schmalbandrauschen für sieben Probanden.

Post-hoc wird der Wilcoxon-Test paarweise (exakt, zweiseitig) angewendet, um die Stimuli aufzufinden, die sich unterscheiden. Ohne Berücksichtigung der α -Fehler Inflation liegt ein signifikanter Unterschied zwischen dem Stimulus mit der Bandbreite 2 Hz und den Stimuli mit 70 Hz ($p=.046$), 140 Hz ($p=.016$) und 270 Hz ($p=.016$) vor. Der Stimulus mit der Bandbreite 10 Hz unterscheidet sich jeweils von 70 Hz ($p=.032$), 140 Hz ($p=.032$) und 270 Hz ($p=.016$). Die Stimuli mit den Bandbreiten 70 Hz und 270 Hz unterscheiden sich ebenfalls signifikant ($p=.032$).

Wird die Bonferroni-Korrektur für die α -Fehler Inflation berücksichtigt, können keine signifikanten Unterschiede zwischen den Bandbreiten ermittelt werden.

6. Einfluss verschiedener Signalparameter auf die Tonhaltigkeit

Eine ANOVA (Anhang F.3.5) zeigt ebenfalls einen Effekt der Bandbreite ($F(4,30)=6.6, p<.001$).

Diskussion

Für geringe Bandbreiten (2 Hz und 10 Hz) ist die THK vergleichbar zur THK von Sinustönen. Die Tonhöhe ist identisch, der Pegel schwankt jedoch leicht aufgrund von Eigenmodulationen des schmalbandigen Rauschens.

Der tonale Eindruck und damit die THK nehmen mit zunehmender Bandbreite monoton ab. Dieser Effekt ist in der DIN 45681:2006-08 modelliert, wie von Sagemühl und Schmidt [SS03] vorgeschlagen. Die Bandbreite des Rauschens muss mit einer Ausprägtheit der Tonhöhe von mindestens 70% relativ zu einem Sinuston vorliegen, um berücksichtigt zu werden, siehe Abschnitt 2.6.

Anhand einer Studie von Fastl [Fas97] wird für die vorliegenden Schmalbandrauschen etwa folgende Ausprägtheit der Tonhöhe ermittelt: $\approx 90\%$ (3 Hz), $\approx 85\%$ (10 Hz), $\approx 60\%$ (70 Hz), $\approx 40\%$ (140 Hz) und $\approx 20\%$ (270 Hz). Diese Werte gelten für Schmalbandrauschen mit konstantem Pegel.

Fruhmann [Fru04] findet für die gleichen Signale bei konstanter Lautheit tendenziell größere Werte für die Ausprägtheit der Tonhöhe: 100% (3 Hz und 10 Hz), $\approx 80\%$ (70 Hz), $\approx 65\%$ (140 Hz) und $\approx 40\%$ (270 Hz).

In beiden Fällen gelten die Angaben der Ausprägtheit für die Schmalbandrauschen ohne weitere überlagerte Rauschen oder andere Signalanteile. Die angegebenen Werte sind für die vorliegende Studie daher als Obergrenze anzusehen, da die Schmalbandrauschen hier in einem Hintergrundrauschen enthalten sind und die Maskierung des Hintergrundes auf das untersuchte Schmalbandrauschen für eine Verringerung der Ausprägtheit sorgt.

Trotzdem scheint die Begrenzung nach DIN 45681:2006-08 auf 75% der Ausprägtheit eines reinen Tones nicht auszureichen. In dieser Studie ruft das Schmalbandrauschen mit einer Bandbreite von 270 Hz eine THK hervor, die etwa 4 dB kleiner ist als die eines reinen Tones, obwohl das Schmalbandrauschen nur eine Ausprägtheit der Tonhöhe von $\approx 20\%$ (Fastl) bzw. $\approx 40\%$ (Fruhmann) besitzt.

Diese These wird durch Gleichung 2.2 der Norm bestätigt: Für die maximal zulässige Bandbreite Δf_R ergibt sich für 700 Hz ein Wert von 44 Hz. Unterhalb der Bandbreite von 44 Hz ruft das Schmalbandrauschen die gleiche THK hervor wie ein reiner Ton, oberhalb von 44 Hz ist die THK der untersuchten Stimuli aber zu groß, um sie zu vernachlässigen.

Im Sinne einer konservativen Abschätzung der THK von Schmalbandrauschen sollte die Norm dies berücksichtigen und die THK im Zweifel eher über- als unterschätzen.

Versuche dieser Art wurden anstelle des hier verwendeten „gaußschen“ Rauschens auch mit Low-Noise Noise durchgeführt (Vormann et al. [VVMS00] und

[VVM⁺99]). Im wesentlichen zeigen die Daten die selben Tendenzen wie die hier diskutierten.

6.4.6. THK als Funktion der Modulationsfrequenz (AM) modulierter Töne

Die DIN 45681:2006-08 unterscheidet nicht zwischen stationären und instationären Geräuschen, d.h. das angegebene Berechnungsverfahren für die THK wird immer auf die gleiche Weise durchgeführt. Es obliegt dem Benutzer, die Voraussetzungen der Anwendbarkeit (bezogen auf die Stationarität der Signale) zu prüfen und das Verfahren entweder nicht anzuwenden oder die Ergebnisse entsprechend selektiv zu interpretieren.

Die Vorgabe der Norm bezüglich schwankender Signalanteile lautet: „Signale mit sehr hoher Pegel- und/oder Frequenzdynamik, die mit einer 3-Sekunden-Mittelung nicht mehr korrespondieren, können durch diese Norm daher nicht mehr beurteilt werden.“ (DIN 45681:2006-08 [THK06], siehe auch Abschnitt 2.10).

Der Einfluss von schwankenden tonalen Komponenten auf die THK wird zunächst mittels einer Amplitudenmodulation (AM) des Tones im Teststimulus ($f_T=700$ Hz) untersucht.

Die Dauer der Stimuli beträgt wie zuvor 1 s. Aus diesem Grund werden zum einen nur Modulationsfrequenzen ≥ 2 Hz betrachtet. Zum anderen wird das von der DIN 45681:2006-08 vorgegebene Mittelungsintervall von 3 s in dieser Untersuchung nicht berücksichtigt.

Eine Modulation f_{mod} von 2 Hz liegt im Bereich der Schwankungsstärke, d.h. ein Zuhörer kann die Schwankung der Frequenz deutlich auflösen und ihr folgen [FZ07]. Die Modulationsfrequenz $f_{mod}=30$ Hz ruft im Wesentlichen eine R-Rauhigkeit hervor [SW94], die Modulationsfrequenz 10 Hz liegt etwa im Übergangsbereich zwischen Schwankungsstärke und R-Rauhigkeit. Eine Modulationsfrequenz von 70 Hz befindet sich nahe dem Bereich, der maximales Rauigkeitsempfinden verursacht [Dan95].

Die Amplitudenmodulation erfolgt mit einem Modulationsgrad von $m=1$ ⁵⁶, die Pegel sind so wie zuvor (Ton 45 dB SPL, GAR 50 dB SPL) gewählt.

Ergebnisse

Die gemessenen Daten sind in Abbildung 6.9 zusammengefasst. Der Friedman-Test ($\chi^2(4, N=7)=11.3$) offenbart eine signifikante ($p=.016$) Abhängigkeit der THK von der Amplitudenmodulation. Die Übereinstimmung der Probanden ist $W=.40$.

⁵⁶ Die Software zur Versuchssteuerung SI [PK96] bietet zur Erzeugung der modulierten Töne den Befehl AMPMOD.

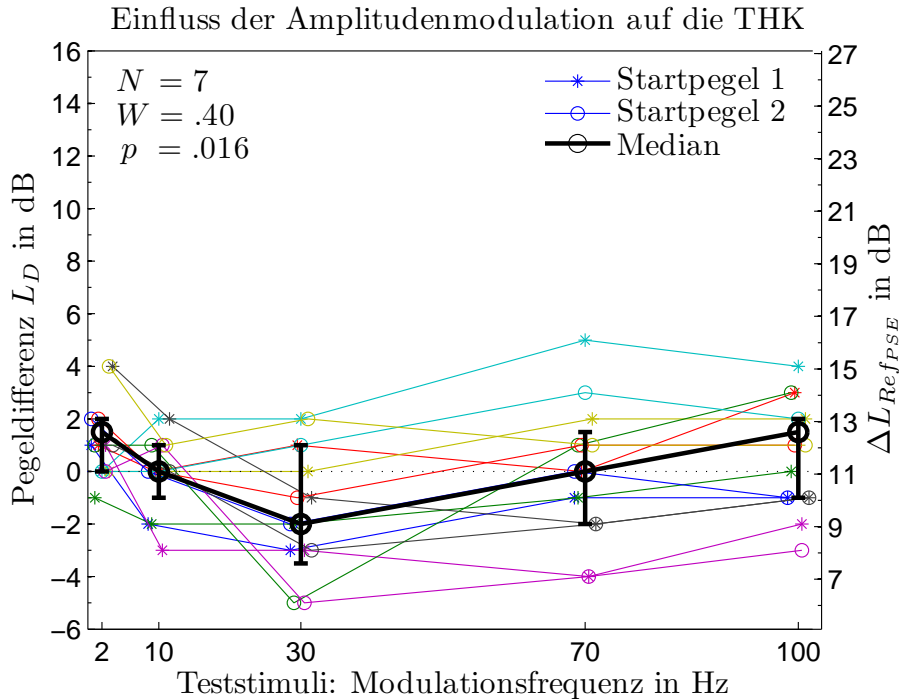


Abbildung 6.9.: Aufgetragen ist die THK (vgl. Abschnitt 6.3.3) über die Modulationsfrequenz der Amplitudenmodulation für sieben Probanden.

Durch paarweise Anwendung des Wilcoxon-Tests lassen sich einzig zwei Unterschiede zwischen den Parametern nachweisen: Die Stimuli mit Modulationsfrequenzen 2 Hz und 10 Hz unterscheiden sich signifikant (unkorrigiert, exakt berechnet, zweiseitig) mit $p=.047$, ebenso wie die Stimuli mit 30 Hz und 100 Hz ($p=.016$). Bei Berücksichtigung der α -Fehler Inflation (Bonferroni) wird kein Unterschied aufgeklärt.

Unter der Annahme intervallskalierter Daten wird zusätzlich eine ANOVA (Anhang F.3.6) durchgeführt. Auch sie belegt eine Abhängigkeit der THK von der Modulationsfrequenz: $F(4,24)=3.24$, $p=.029$.

Diskussion

Der schwankende Ton ($f_{mod}=2$ Hz) zeigt eine größere THK ($L_D=1,5$ dB, Interquartilbereich 0 dB bis 2 dB) als es für einen nicht modulierten Ton zu erwarten wäre. Das kann damit erklärt werden, dass der Ton aufgrund seiner steten Veränderung auffälliger ist als ein stationärer Sinuston.

Für die anderen Modulationsfrequenzen beinhalten die Interquartilbereiche der gemessenen THK die Pegeldifferenz $L_D=0$ dB. Diese Modulationsfrequenzen scheinen daher keinen wesentlichen Einfluss auf die THK zu haben, auch wenn

der Friedman-Test einen signifikanten Effekt anzeigt. Auch die ANOVA-Analyse (Anhang F.3.6) zeigt, dass die Bewertung der THK nicht erwähnenswert von der Amplitudenmodulation abhängt. Wenn überhaupt, gibt es lediglich einen Effekt im Bereich der Schwankungsstärke.

Anhand dieser Daten kann die Vorgehensweise der DIN 45681:2006-08 als sinnvoll bestätigt werden. Falls Amplitudenmodulationen einen Effekt auf die THK haben ist dieser klein und kann daher vernachlässigt werden.

6.4.7. THK als Funktion der Modulationsfrequenz (FM) modulierter Töne

Die Untersuchung zur Modulationsfrequenz (AM) modulierter Töne wird im Folgenden für Frequenzmodulationen (FM) wiederholt, die Pegel (Ton 45 dB SPL, GAR 50 dB SPL) und die Modulationsfrequenzen werden wie zuvor festgelegt.

Der Frequenzhub bei der Frequenzmodulation beträgt für alle Stimuli 100 Hz, d.h. die Töne schwanken periodisch zwischen 600 Hz und 800 Hz.

Ergebnisse

Die Ergebnisse sind in Abbildung 6.10 zusammengefasst. Der Friedman-Test ($\chi^2(4, N=7)=11.2$) deckt eine signifikante Abhängigkeit der THK von der Modulationsfrequenz (FM) auf. Die Signifikanz ($p=.017$) und die Übereinstimmung der Probanden ($W=.40$) ist vergleichbar zu den Werten für die Amplitudenmodulation.

Durch paarweise Anwendung des Wilcoxon-Tests lässt sich lediglich ein Unterschied zwischen den Parametern finden. Die beiden Stimuli mit einer Modulationsfrequenz von 10 Hz bzw. 70 Hz unterscheiden sich signifikant ($p=.031$, unkorrigiert, exakt berechnet, zweiseitig). Bei Berücksichtigung der α -Fehler Inflation (Bonferroni) kann kein signifikanter Unterschied auf dem 5%-Niveau identifiziert werden.

Unter der Annahme intervallskalierter Daten wird eine ANOVA ($F(4,24)=4.2$, $p=.010$) durchgeführt, vgl. Anhang F.3.7. Diese bestätigt die hier gefundenen Ergebnisse, liefert aber keine wesentlich neuen Erkenntnisse.

Diskussion

Auch für Frequenzmodulationen scheint es die Tendenz zu geben, dass für kleine Modulationsfrequenzen die THK leicht angehoben wird. Dieser Effekt ist jedoch nicht statistisch eindeutig nachzuweisen.

Ansonsten sind die Daten für die Frequenzmodulation ähnlich zu den Daten der Amplitudenmodulation. Daher folgen die gleichen Bewertungen und Schlüsse, insbesondere dass die DIN 45681:2006-08 eine zulässige Vereinfachung

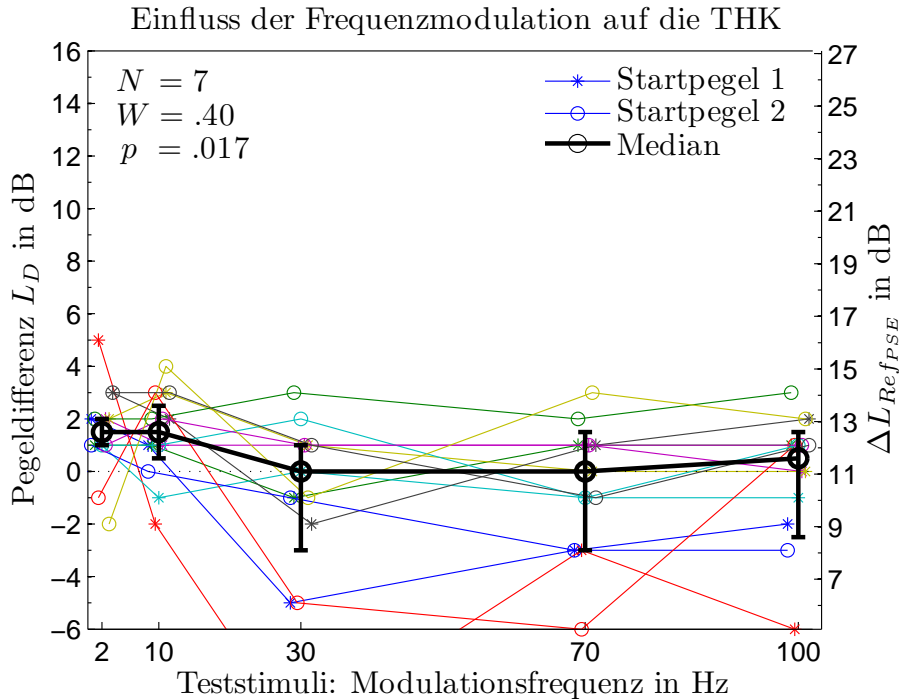


Abbildung 6.10.: Aufgetragen ist die THK (vgl. Abschnitt 6.3.3) über die Modulationsfrequenz (FM) für sieben Probanden.

vornimmt, indem sie lediglich die gemittelten Spektren für die Berechnung der maßgebliche Differenz heranzieht.

6.4.8. THK als Funktion der Tonfrequenz

Sinustöne unterschiedlicher Frequenzen aber konstantem Pegel von 45 dB SPL werden in GAR bei 50 dB SPL bzgl. ihrer THK untersucht. Die Frequenzen sind auf einer Bark-Skala gleichabständig (1,5 Bark) angeordnet, beginnend bei 2 Bark und endend mit 17 Bark.

Ergebnisse

Die Daten dieser Untersuchung sind in Abbildung 6.11 zusammengefasst.

Der Friedman-Test ($\chi^2(10, N=8)=34.4$) bestätigt den visuellen Eindruck aus Abbildung 6.11. Die Frequenz hat einen signifikanten Einfluss ($p < .001$) auf die THK.

Paarweise durchgeführte Wilcoxon-Tests zeigen, dass sich der Ton mit 200 Hz von den Tönen ≥ 510 Hz signifikant ($.008 \leq p < .023$ zweiseitig, exakt berechnet, unkorrigiert) unterscheidet. Der Ton mit 350 Hz unterscheidet sich ebenfalls signifikant ($.008 \leq p < .047$.) von den Tönen ≥ 510 Hz.

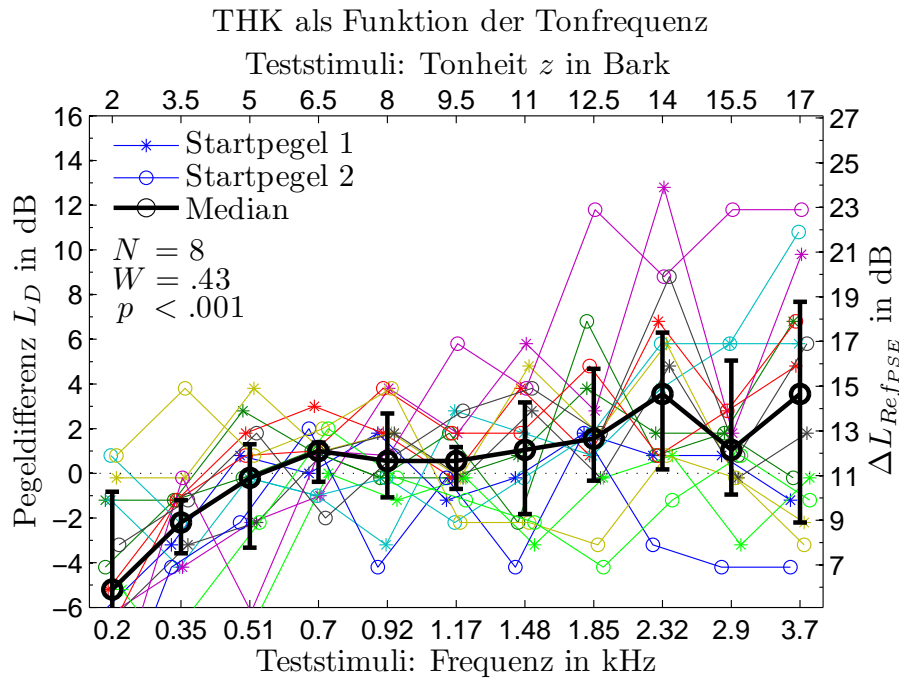


Abbildung 6.11.: Einfluss der Tonfrequenz auf die THK (vgl. Abschnitt 6.3.3) für acht Probanden. Im unteren Bereich sind die Tonfrequenzen in Hertz, im oberen Bereich die Tonheit in Bark angegeben.

Bei Berücksichtigung der α -Fehler Inflation (Bonferroni) können keine signifikanten Unterschiede nachgewiesen werden.

Da sich nur die beiden kleinsten Frequenzen deutlich von den anderen unterscheiden, liegt die Übereinstimmung (Kendall- W) der Probanden in einem eher moderaten Bereich von $W=.43$.

Unter der Annahme intervallskalierter Ergebnisse wird eine ANOVA (siehe Anhang F.3.8) durchgeführt. Auch sie zeigt einen signifikanten Effekt der Frequenz ($F(10,70)=6.3$, $p<.001$) auf die THK. Post-hoc durchgeführte paarweise T -Tests (Bonferroni) zeigen im Wesentlichen die gleichen Unterschiede wie die nonparametrischen Verfahren, siehe Tabelle F.9 im Anhang.

Wird die ANOVA ohne die beiden Frequenzen 200 Hz und 350 Hz wiederholt, so ist ein Effekt der Frequenz auf die THK auf dem 5%-Signifikanzniveau nicht mehr nachweisbar ($F(8,56)=1.7$, $p=.12$).

Das gilt auch für die parameterfreie Analyse: Der Friedman-Test, angewendet auf den Datensatz ohne die beiden Stimuli mit der kleinsten Frequenz (200 Hz und 350 Hz), zeigt keinen Einfluss der Frequenz ($\chi^2(8,N=8)=14.3$, $p=.074$) auf die THK.

Diskussion

Bei der Analyse der Daten fällt auf, dass der Interquartilbereich für die acht Probanden größer wird, je weiter die untersuchte Frequenz der Teststimuli von der Frequenz des Referenztones (700 Hz) entfernt liegt. Das ist nachvollziehbar, da sich die tonale Wahrnehmung der verschiedenen Frequenzen für größere Frequenzdifferenzen stärker ändert und die Probanden dadurch unsicherer werden.

Aus diesem Grund kann nicht ohne Weiteres geklärt werden, inwieweit der gefundene Frequenzeinfluss durch die Wahl des Referenztones beeinflusst wird, da es sich auch um einen generellen Effekt handeln könnte, der nicht von der Frequenz des Referenztones abhängig ist.

Aus diesem Grund wird die Messung zunächst mit einem alternativen Referenzton wiederholt, u.a. um die Reproduzierbarkeit der Ergebnisse zu verifizieren.

Die Diskussion der Ergebnisse an sich wird daher erst im Anschluss an die nächste Untersuchung geführt.

6.4.9. THK als Funktion der Tonfrequenz II

Es werden die gleichen Stimuli wie in der vorigen Fragestellung bewertet, jedoch mit einem alternativen Referenzton als Vergleichsstimulus. Die Frequenz des Referenztones $f_{T_{Ref}}$ wird hier zu 1850 Hz gewählt, d.h. bezogen auf die Teststimuli, symmetrisch zu höheren Frequenzen gespiegelt.

Ergebnisse

Die gemessenen Daten zeigt Abbildung 6.12. Die linke Ordinate ist aufgetragen wie bisher, hat aber für diese Messung einen anderen Bezugspunkt. Aus diesem Grund muss die rechte Ordinate gegenüber Abbildung 6.11 angepasst werden. Analog zu Gleichung 6.13 ergibt sich ΔL_{RefPSE} mit dem Verdeckungsmaß $a_v(1850 \text{ Hz}) = -3,4 \text{ dB}$ an dieser Stelle zu

$$(6.15) \quad \Delta L_{RefPSE} = L_D + 12,0 \text{ dB}.$$

Der Trend für diese Messung ist nicht so eindeutig zu erkennen wie in der Untersuchung zuvor. Trotzdem bestätigt der Friedman-Test ($\chi^2(10, N=8) = 40.9$) den signifikanten Einfluss ($p < .001$) der Frequenz auf die THK.

Auch in dieser Messung haben die Stimuli mit kleineren Frequenzen den größten Anteil an diesem Effekt. Das zeigen paarweise durchgeführte Wilcoxon-Tests: Der Stimulus mit 200 Hz unterscheidet sich signifikant ($.008 \leq p < .016$ zweiseitig, exakt berechnet, unkorrigiert) von allen anderen.

Der Stimulus mit 350 Hz unterscheidet sich signifikant von den folgenden Stimuli: 700 Hz ($p = .016$), 920 Hz ($p = .008$), 1850 Hz ($p = .008$), 2320 Hz ($p = .016$) und 2900 Hz ($p = .039$).

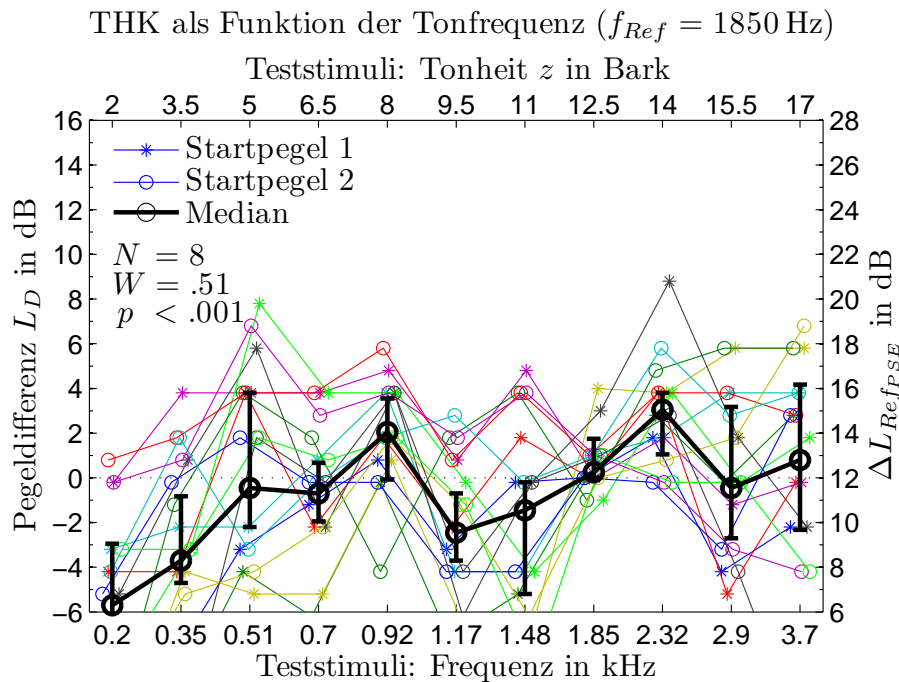


Abbildung 6.12.: Einfluss der Tonfrequenz auf die THK für den Referenzton mit einer Frequenz $f_{T_{Ref}} = 1850$ Hz. Im Vgl. zu Abbildung 6.11 muss die rechte Ordinate aufgrund der unterschiedlichen Frequenzen der Referenzöne hier angepasst werden, siehe Text.

Für diese Messung existieren fünf weitere signifikante (unkorrigierte) Unterschiede zwischen den Stimuli: 920 Hz und 1170 Hz ($p = .008$), 920 Hz und 1480 Hz ($p = .047$), 1170 Hz und 1850 Hz ($p = .008$), 1170 Hz und 2320 Hz ($p = .008$) sowie 2320 Hz und 2900 Hz ($p = .047$).

Die beiden Frequenzen 1170 Hz und 1480 Hz haben, verglichen mit den Ergebnissen der vorigen Untersuchung, verhältnismäßig kleine THK-Werte. Aus diesem Grund wird der Effekt der Frequenz auf die THK nicht ausschließlich von den kleinsten Frequenzen bestimmt, sondern zusätzlich auch von diesen beiden Frequenzen.

Bei Berücksichtigung der α -Fehler Inflation (Bonferroni) können keine signifikanten Unterschiede auf dem 5%-Niveau nachgewiesen werden.

Die Übereinstimmung der Probanden (Kendall-W) beträgt hier $W = .51$ und liegt damit in einem ähnlichen Bereich wie bei der Untersuchung zuvor.

Unter der Annahme intervallskalierter Ergebnisse wird eine ANOVA durchgeführt. Sie zeigt einen signifikanten Effekt der Frequenz ($F(10,77) = 6.6$, $p < .001$), vgl. Anhang F.3.8.

Diskussion

Die beiden in den Abbildungen 6.11 und 6.12 visualisierten Datensätze werden gemeinsam ausgewertet und diskutiert. Da sie für verschiedene Frequenzen des Referenztones gemessen werden, muss die absolute Skalierung der THK, d.h. in diesem Fall ΔL_{RefPSE} (die rechten Ordinaten der Abbildungen 6.11 und 6.12) verwendet werden, um die Daten unmittelbar vergleichen zu können. Das zeigt Abbildung 6.13 für die beiden Datensätze.

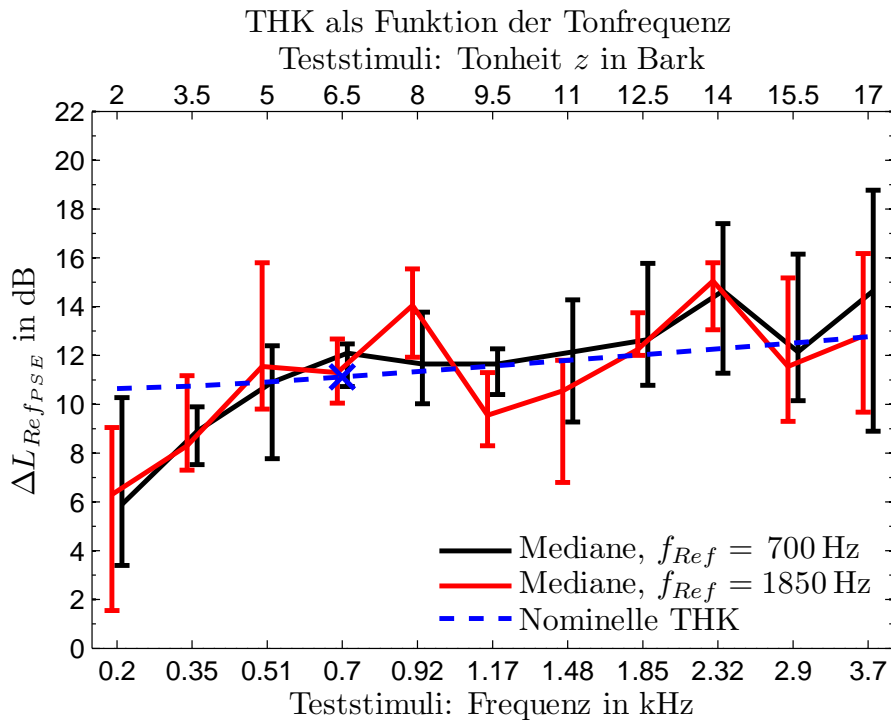


Abbildung 6.13.: Gemessene und nominelle THK über den Tonfrequenzen der Teststimuli für acht Probanden, siehe Text.

Es ist abzulesen, dass die beiden Messungen gut übereinstimmen. Lediglich für die beiden Frequenzen 1170 Hz und 1480 Hz liegt der Median der einen Messung deutlich außerhalb des Interquartilbereiches der anderen Messung. Für 920 Hz liegen die Interquartilbereiche der einen Messung in der Nähe des Medians der anderen Messung.

Anhand von Gleichung 2.1 wird der nominelle Wert der THK für die elf Teststimuli ermittelt:

$$(6.16) \quad \Delta L_{Test} = L_{T_{Test}} - L_{GAR} - a_v(f_{T_{Test}}).$$

Dieser Wert ist aufgrund des Verdeckungsmaßes frequenzabhängig. In Abbildung 6.13 ist er blau gestrichelt dargestellt.

Für Frequenzen ≥ 510 Hz zeigen die gemessenen Daten relativ genau den Verlauf der nominellen, d.h. der durch die DIN 45681:2006-08 vorhergesagten THK. Der Ansatz der DIN 45681:2006-08, die Frequenzabhängigkeit der THK über das Verdeckungsmaß zu berücksichtigen, ist für diesen Frequenzbereich somit adäquat.

Ein ähnlicher Verlauf der THK über der Frequenz wurde bereits bei den Untersuchungen mittels Paarvergleichen im Kapitel 5.6, Abbildung 5.7, gefunden. Aus Abbildung 6.13 kann darüber hinaus der Unterschied zwischen der nach DIN 45681:2006-08 vorhergesagten und der tatsächlich wahrgenommenen THK ermittelt werden.

Für Frequenzen unterhalb von 510 Hz wird die THK überschätzt. Für 200 Hz liegt die vorhergesagte (nominelle) THK etwa 4–5 dB über der gemessenen THK, für 350 Hz ist die Differenz etwa 2 dB.

Hellweg und Nobile [HN02] präsentieren ähnliche Ergebnisse für Frequenzen um 250 Hz, obwohl andere Stimuli verwendet werden und der Ansatz zur Bewertung sich erheblich von dem hier verwendeten unterscheidet.

Wie Hansen [Han10] zeigt, hängt die THK auch von der Ausgeprägtheit der Töne („Pitch Strength“) ab. Die Ausgeprägtheit der Töne in den verwendeten Teststimuli kann nach Fastl und Zwicker ([FZ07], Abbildung 5.28) abgeschätzt werden. Zwischen 750 Hz und 3000 Hz ist die Ausprägung konstant ($\approx 100\%$), während sie zu größeren und kleineren Frequenzen hin deutlich abfällt.

Der Testton mit 510 Hz hat eine Ausgeprägtheit von $\approx 80\%$ relativ zum Bereich maximaler Ausgeprägtheit. Der Testton mit 350 Hz weist einen etwas geringeren Wert von $\approx 70\%$ auf. Die Ausgeprägtheit des Testtons mit 200 Hz wird zu $\approx 35\%$ ermittelt.

Anhand der gemessenen Daten kann die THK-Abnahme aufgrund der abnehmenden Ausgeprägtheit der Töne quantifiziert werden. Dazu ist lediglich die Differenz der nominellen und der gemessenen THK aus Abbildung 6.13 zu entnehmen.

Im Kapitel 5.6 zeigt sich, dass es zwei verschiedene Gruppen von Probanden gibt, die die THK bzgl. der Frequenzabhängigkeit unterschiedlich bewerten. Dieser Effekt ist auch hier zu beobachten. In Abbildung 6.11 bewerten zwei Probanden (dunkelblau und gelbgrün) die THK für Frequenzen oberhalb von 2320 Hz abfallend für anwachsende Frequenzen. In Abbildung 6.12 trifft das für drei Probanden (schwarz, lila, hellgrün) zu.

Einer der insgesamt fünf Probanden mit diesem Bewertungskriterium hat an beiden Studien (Referenzton 700 Hz bzw. 1850 Hz) teilgenommen. Die anderen drei Probanden jeweils nur an einer der beiden Messungen. Insgesamt haben fünf Probanden beide Messungen durchgeführt, jeweils drei weitere Probanden nur eine der beiden Messungen.

Es zeigt sich somit, dass die Ergebnisse aus Kapitel 5.6 allesamt bestätigt werden können. Aufgrund des quantitativen Messverfahrens können die Unter-

schiede mit dem hier benutzten Messverfahren direkt erfasst und beschrieben werden.

6.5. Zusammenfassung

Anhand der zuvor präsentierten und diskutierten Ergebnisse zeigt sich die Mächtigkeit des vorgestellten Verfahrens. Die gefundenen Zusammenhänge können allesamt sinnvoll interpretiert und in einen größeren Zusammenhang mit anderen psychoakustischen Befunden integriert werden.

Trotz einer vergleichsweise geringen Anzahl von Probanden ist die Aussagekraft der meisten gefundenen Ergebnisse statistisch belegbar.

Die Ergebnisse des genutzten adaptiven Verfahrens zur Bewertung der THK sind Pegelwerte, die unmittelbar zu interpretieren sind und quantitative Aussagen ermöglichen.

Die für die vorliegenden Untersuchungen ausgewählten Startpegel haben keinen nachweisbaren Einfluss auf die Messergebnisse.

Offenbar ist das propagierte Verfahren robust gegen die meisten typischen Fehler bei AFC-Messungen. Um diese These genauer zu untersuchen und um abzuschätzen, ob und wie der Aufwand für das AFC-Verfahren reduziert werden kann, wird diese Fragestellung im nächsten Kapitel erneut aufgegriffen und untersucht.

Im Hinblick auf die DIN 45681:2006-08 lassen sich anhand der gemessenen Daten die folgenden Aussagen konstituieren. Dabei muss jedoch darauf hingewiesen werden, dass neben der reinen Betrachtung von Effektgrößen und der Berechnung von statistischen Massen stets die praktische Bedeutsamkeit im Vordergrund stehen muss [BD06]. Der weitaus wichtigste Parameter für die Berechnung der THK ist die maßgebliche Differenz ΔL . Diese wird nur unbedeutend durch den absoluten Pegel der Darbietung beeinflusst (Kapitel 6.4.4).

Treten mehrere (Teil-) Töne auf, so kann kein Einfluss der Harmonizität der Teiltöne untereinander festgestellt werden (Kapitel 6.4.2). Abgesehen von virtuellen Tonhöhen, die hier nicht untersucht werden, ist lediglich die Anzahl der (Teil-) Töne (Kapitel 6.4.1) und ihr Abstand zueinander (Kapitel 6.4.3) entscheidend, siehe nächsten Abschnitt.

Instationäre Signale wie die untersuchte AM (Kapitel 6.4.6) und FM (Kapitel 6.4.7) sind für die Beurteilung der THK nicht bedeutend, solange der vorgeschriebene Algorithmus in der DIN 45681:2006-08 in der Lage ist, die korrekten Tonpegel zu ermitteln. Die untersuchten Modulationen selber haben keinen Einfluss auf die wahrgenommene THK. Deshalb muss diese Art der Instationarität nicht besonders in der Norm berücksichtigt werden.

Wie die THK über längere Zeitabschnitte integriert wird, kann daraus allerdings nicht geschlussfolgert werden. Diese Fragestellung ist Gegenstand aktueller Forschung.

Die Frequenzabhängigkeit der THK (Kapitel 6.4.8 und 6.4.9) wird in der DIN 45681:2006-08 für Frequenzen oberhalb von 500 Hz angemessen durch die Mithörschwelle, bzw. daraus resultierend das Verdeckungsmaß, berücksichtigt.

Unterhalb von 500 Hz wird die THK durch diese Art der Modellierung jedoch überschätzt. Im Rahmen des Anwendungsbereiches der Norm ist dieses Verhalten zu tolerieren, da es im Zweifel die tatsächliche Belästigung durch die tonale Komponente überschätzt und daher Geräuschmissionen eher konservativ bewertet.

Die größte Schwäche der Norm besteht darin, für zwei oder mehr tonale Komponenten mit vergleichbar großer maßgeblicher Differenz ΔL (Kapitel 6.4.1), die Integration dieser Komponenten nicht angemessen zu berücksichtigen. Die Norm postuliert, nur die Komponente mit der größten maßgeblichen Differenz zu betrachten, während die vorliegenden Ergebnisse eher darauf hinweisen, dass alle Komponenten gemeinsam betrachtet werden müssen. Dieser Zusammenhang wird im nächsten Kapitel weiter untersucht.

Hinzu kommt, dass der Abstand zweier tonaler Komponenten (Kapitel 6.4.3) nicht ausschließlich dadurch berücksichtigt werden kann, ihre Energie für kleine Abstände der beiden Komponenten aufzuaddieren. Vielmehr wächst die THK für die untersuchten Abstände über die reine Addition der Energien (+3 dB) hinaus an.

Auch Schmalbandrauschen (Kapitel 6.4.5) werden tendenziell von der Norm unterschätzt. Für die untersuchten Bandbreiten unterhalb von 44 Hz wird die THK korrekt abgebildet, während die Norm für größere Bandbreiten einen Effekt auf die THK schlicht nicht berücksichtigt. Dieser Effekt ist kleiner als die zuvor diskutierten Effekte (besonders die Abhängigkeit von der Anzahl etwa gleich großer tonaler Komponenten) aber durchaus in einem Bereich, der noch berücksichtigt werden sollte.

Der Berechnungsalgorithmus der DIN 45681:2006-08 ist vorgesehen, um A-bewertete Spektren zu verarbeiten (Kapitel 2.2). Allerdings würde die A-Bewertung die Eleganz des vorliegenden Bewertungsverfahrens und der durch die Verwendung des GAR äußerst einfachen theoretischen Beschreibung der Frequenzgruppenpegel beeinträchtigen. Aus diesem Grund wird die A-Bewertung zugunsten einer leicht nachvollziehbaren Darstellung vernachlässigt. Die Vernachlässigung stellt keine ernsthafte Einschränkung dar, weil die A-Bewertung der Spektren bezüglich der maßgeblichen Differenzen nur einen kleinen Einfluss auf die Berechnung der THK hat.

Alle Untersuchungen werden bei einer nominellen maßgeblichen Differenz von $\Delta L=11,1$ dB durchgeführt, bis auf die Messung mit dem Referenzton bei

1850 Hz. 11,1 dB kennzeichnet eine eher ausgeprägte THK, was aus dem Tonzuschlag von $K_T=5$ dB für diese maßgebliche Differenz abzulesen ist.

Für die vorliegenden Untersuchungen wird die maßgebliche Differenz bewusst hoch angesetzt. So besteht die Möglichkeit, die wahrgenommene THK für den Teststimulus durch Verringerung der maßgeblichen Differenz des Vergleichsstimulus (d.h. den Referenzstimulus) um den entsprechenden Pegelwert der maßgeblichen Differenz (11,1 dB) weniger tonhaltig zu bewerten. Wird die maßgebliche Differenz zu klein gewählt, so können einige Effekte nicht korrekt vorhergesagt werden, da für einige Probanden die individuellen THK-Daten dann unter die Mithörschwelle fallen müssten, vgl z.B. Abbildung 6.8 oder 6.12.

Nachdem die Größenordnung der einzelnen Effekte bekannt ist, wird es möglich, die THK in einem eher schwelennahen Bereich zu untersuchen. Andererseits läßt die einschlägige Literatur zum Thema THK vermuten, dass die gefundenen Abhängigkeiten der THK ebenso für kleinere und größere maßgebliche Differenzen gültig ist. Streng genommen müsste diese Annahme allerdings sowohl für größere als auch für kleinere maßgebliche Differenzen überprüft werden.

7. Verfahrensvergleich zur Bestimmung der Tonhaltigkeit

Im voran gegangenen Kapitel ist das adaptive THK-Messverfahren vorgestellt und seine Leistungsfähigkeit demonstriert worden. Im Folgenden wird der Einfluss der Konfiguration des Verfahrens, d.h. die Wahl der Parameter und des Messdesigns, auf die Messung der THK ermittelt.

Dazu werden zunächst Untersuchungen des Einflusses von Messparametern für Pegel gleicher Lautheit diskutiert, da sie ein vergleichbares Messverfahren benutzen wie das hier propagierte. Die im Folgenden zitierten Studien zeigen ungewollte Einflüsse (sogenannte Bias-Effekte) auf, die bei der Messung der Tonhaltigkeit minimiert werden sollen.

Verhey [Ver99] zeigt für die Messung des Pegels gleicher Lautheit von Bandpassrauschsignalen, dass im Gegensatz zu einfachen, nacheinander durchgeführten adaptiven Messungen (Tracks) die Ergebnisse von interleaved Messungen⁵⁷ nicht vom Startpegel der Tracks abhängen. Verhey [Ver99] und Verhey und Uhlemann [VU08] finden weiterhin, dass die Wahl der Referenz (in dieser Studie die Referenzbandbreite) im Mittel keinen Einfluss auf die mittleren Pegel gleicher Lautheit hat. Für Einzelpersonen kann das Ergebnis der Messung hingegen von der Wahl der Referenz abhängen.

Buus et al. [BFP97] zeigen ebenfalls eine Abhängigkeit der Ergebnisse von den experimentellen Parametern. Hier werden für zwei sich leicht überlappende Pegelbereiche die Pegel gleicher Lautheit für kurze und lange Signale bestimmt. Es zeigen sich Unterschiede der Ergebnisse von bis zu etwa 10 dB im Überlappungsbereich. Außerdem unterscheiden sich die Daten, je nachdem ob das lange oder das kurze Signal in der adaptiven Messung variiert wird. Beides deutet darauf hin, dass auch bei interleaved Messungen das Messdesign (d.h. das experimentelle Paradigma) die Messergebnisse beeinflussen kann.

Für eine effektive Anwendung des AFC-Verfahrens muss die Schrittweite während der Messphase zu den Reizen passen. Eine Schrittweite unterhalb des JND⁵⁸

⁵⁷ Interleaved-Verfahren: die simultane Messung mehrerer Parameter, wobei die Tracks für die einzelnen Parameter ineinander verschachtelt sind

⁵⁸ JND = „Just noticeable Difference“, d.h. der eben merkbare Unterschied zwischen zwei Ausprägungen einer Empfindungsgröße. Im Englischen auch als „Difference Limen“ bezeichnet.

verlängert die Messung unnötig, während eine minimale Schrittweite deutlich oberhalb des JNDs die Messung zu Lasten der Genauigkeit verkürzt. Für eine effektive Durchführung ist die Wahl einer adäquaten minimalen Schrittweite daher wichtig.

Als Kompromiss zwischen Aufwand und Genauigkeit wird im Folgenden eine minimale Schrittweite von 2 dB gewählt. Dieser Wert führt in Kapitel 6.3.2 zu sinnvollen Ergebnissen und wird daher beibehalten.

Als Testsignale für die folgenden Messungen werden harmonisch komplexe Töne (HKT) analog zu Kapitel 6.4.1 verwendet. Einerseits ist durch diese Wahl sicher gestellt, dass ein relativ großer Effekt existiert und die unterschiedlichen Messparadigmen und -konfigurationen somit sinnvoll gegeneinander verglichen werden können. Zum anderen wird die Datenlage bzgl. der THK für die HKTs deutlich verbessert und die Ergebnisse so statistisch besser abgesichert.

Es zeigt sich bei der Untersuchung der Frequenzabhängigkeit der THK (Kapitel 6.4.9), dass die THK durch die Mithörschwelle der verwendeten Töne modelliert werden kann. Aus diesem Grund wird im Folgenden zunächst die Mithörschwelle für die verwendeten HKTs ermittelt. Auf diese Weise kann überprüft werden, ob es ebenfalls möglich ist, die THK für HKTs anhand der Mithörschwelle vorherzusagen bzw. zu modellieren.

7.1. Stimuli und Messverfahren

Die Probanden haben die Aufgabe, die THK mittels des in Kapitel 6.2 beschriebenen adaptiven Verfahrens zu bewerten. Dabei wird das THK-Referenzsignal aus Kapitel 6.2.2 verwendet. Die Probanden werden vor den Bewertungen anhand der Standard-Orientierung, siehe Kapitel 4.7 unterwiesen⁵⁹. Die Instruktion erfolgt analog wie im Kapitel 6.2.1.

Die Teststimuli sind prinzipiell die gleichen HKTs in GAR wie in Kapitel 6.4.1. Allerdings wird in diesem Fall der Gesamtpegel der HKTs konstant gehalten, d.h. der Pegel der Teiltöne L_l wird bei Verdopplung der Anzahl von Teiltönen um 3 dB abgesenkt. Im Kapitel 6.4.1 ist hingegen der Pegel der Teiltöne konstant, d.h. der Gesamtpegel der HKTs nimmt bei Verdopplung der Anzahl von Teiltönen nach Gleichung 6.14 um 3 dB zu.

Der Pegel der Teiltöne $L_l(k)$ ergibt sich daher aus dem Gesamtpegel des HKTs L_{HKT} als

$$(7.1) \quad L_l(k) = L_{HKT} - 10 \log(k) \text{dB} \quad \text{für } l = 1 \dots k.$$

Diesen Zusammenhang für die Teststimuli zeigt Abbildung 7.1. Der Gesamtpegel des Rauschens L_{GAR} beträgt wie zuvor 50 dB SPL. Der HKT($k=1$) hat hier

⁵⁹ Die Orientierung erfolgt nur, falls sie zuvor nicht bereits durchgeführt worden ist.

mit 50 dB SPL einen um 5 dB höheren Pegel als in Kapitel 6.4.1. Auf diese Weise sind die Teiltonpegel $L_l(k)$ der hier verwendeten HKTs im Mittel über alle HKTs in etwa vergleichbar zu den im Kapitel 6.4.1 verwendeten Stimuli. Zusätzlich liegen die Teiltöne für den HKT($k=8$) damit deutlich oberhalb des Frequenzgruppenpegels für das Rauschen, so dass alle HKTs auf jeden Fall eine THK hervorrufen.

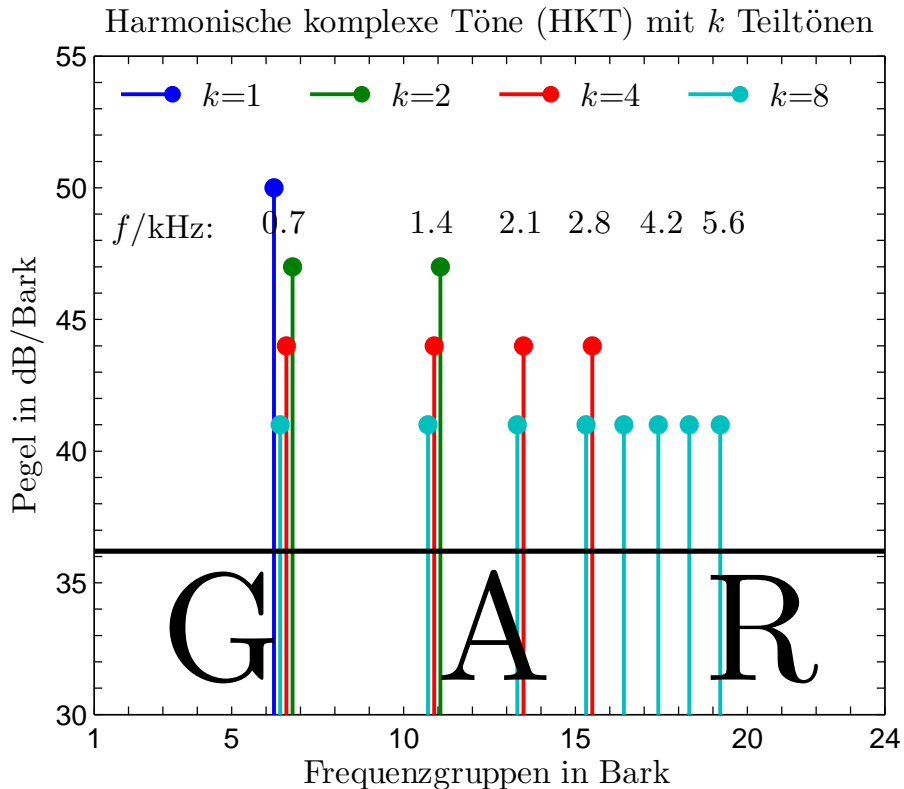


Abbildung 7.1.: Schematische Darstellung der HKTs als Dichtespektrum. Auf der Abszisse sind Frequenzgruppen in Bark aufgetragen. Die Ordinate gibt die Pegel der Teiltöne $L_l(k)$ in dB/Bark an, vgl. Formel 7.1. Der Pegel des GAR wird in dieser Darstellung durch die horizontale Gerade bei 36,2 dB symbolisiert.

7.2. Signaldarbietung

Die Generierung der Stimuli erfolgt während des Experimentes über eine Silicon Graphics (SGI) Workstation mit 16 Bit und einer Abtastfrequenz f_s von 32 kHz. Die Stimuli werden diotisch über einen internen D/A-Wandler (Ariel-Karte) per

Kopfhörer (Sennheiser HDA200) in einer schallisolierten Kabine dargeboten. Die Länge der Signale beträgt 1000 ms, die Pause zwischen zwei Signalen 500 ms.

Da die Abtastfrequenz hier gegenüber Kapitel 6.4 erhöht wird, ergibt sich für den Frequenzgruppenpegel des GAR L_{GAR} nach Gleichung 6.1 verglichen mit Kapitel 6.4.1 ein etwas kleinerer Wert von 36,2 dB.

7.3. Probanden

An den im Folgenden aufgeführten Experimenten nehmen jeweils die gleichen elf⁶⁰ mit psychoakustischen Messungen erfahrenen Personen im Alter zwischen 24 und 31 Jahren (Median 26 Jahre) teil. Sechs Probanden sind weiblich, fünf männlich. Es handelt sich um Mitglieder der Arbeitsgruppe Akustik, Fachbereich Physik und des „Graduiertenkollegs Psychoakustik“ an der Universität Oldenburg. Die Probanden sind normalhörend⁶¹ und werden nicht für ihre Mitarbeit entlohnt.

7.4. Mithörschwelle für die HKTs

7.4.1. Verfahren

Die Mithörschwellen für die vier Teststimuli werden in ausbalancierter Reihenfolge in einzelnen Tracks mit einem 3I-3AFC 1-up 2-down Verfahren gemessen, siehe Kapitel 6.2.1. Bei diesem Verfahren werden jeweils drei Intervalle präsentiert, die aus einem identischen, im Pegel konstanten GAR bestehen. Eines der drei Intervalle enthält zusätzlich den HKT. Die Aufgabe der Probanden besteht darin, das Intervall zu benennen, welches den HKT enthält. Nachdem der Proband sein Urteil abgegeben hat, bekommt er jeweils eine Rückmeldung („Feedback“) darüber, ob seine Antwort richtig war oder nicht.

Mit der verwendeten 1-up 2-down Pegelsteuerung wird der Pegel der HKTs adaptiv verändert und der Wert bei 70,7% der psychometrischen Funktion angestrebt [Lev71]. Die individuelle Wahrscheinlichkeit den HKT vor dem Hintergrund zu detektieren ist daher größer, als den Ton nicht detektieren zu können.

Die Anfangsschrittweite beträgt 8 dB und wird halbiert, bis die Schrittweite noch 1 dB beträgt. In der sich anschließenden Messphase werden zehn Umkehrpunkte gemessen und daraus der Median mit Interquartilbereich ermittelt, der als Messergebnis angegeben wird.

Die Messung der Mithörschwelle dauert für einen HKT etwa 5 Minuten, d.h. pro Proband beträgt die Messdauer für alle vier HKTs im Mittel ca. 20 Minuten.

⁶⁰ Bei der Messung der Mithörschwelle nehmen nur zehn der elf Probanden teil.

⁶¹ Selbsteinschätzung – bei Zweifel an der Normalhörigkeit wird ein Audiogramm aufgenommen.

Die Stimuli (vgl. Abbildung 7.1) werden mit identischen Pegelwerten wie in allen folgenden Messungen präsentiert. Der Gesamtpegel des Rauschens L_{GAR} beträgt 50 dB SPL. Der gleiche Wert wird auch als Startpegel (Gesamtpegel) der verwendeten HKTs L_{HKT} für die adaptive 3I-3AFC-Messung verwendet.

7.4.2. Ergebnisse

Die aggregierte Mithörschwelle $L_M(k)$ für jeden HKT(k) wird durch Bildung des Mittelwerts samt Standardabweichung über die individuellen Probandendaten gebildet. Das Ergebnis ist in Abbildung 7.2 visualisiert.

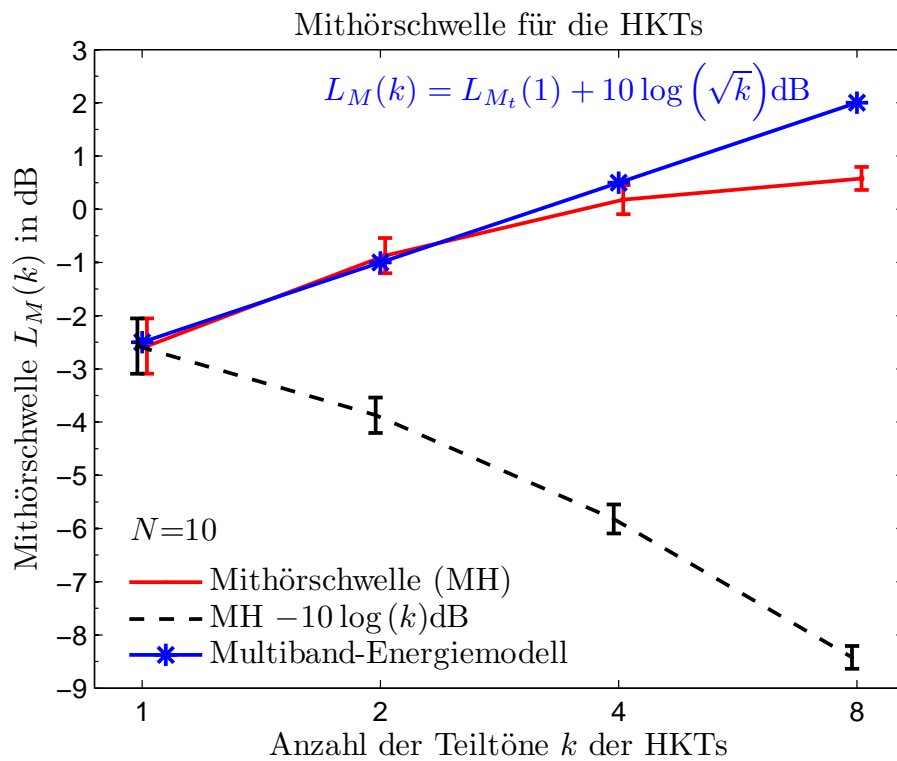


Abbildung 7.2.: Mithörschwelle (Mittelwert und Standardfehler) für die HKTs, aufgetragen analog zum Verdeckungsmaß. Zum Vergleich die Vorhersage des Multiband-Energiemodells und die absolute Maskierungsschwelle für die HKTs bei Vernachlässigung der Pegelkorrekturen aufgrund der unterschiedlichen Anzahl der Teiltöne, siehe Text.

Die Urteiler-Konkordanz (Kendalls W) zwischen den Probanden ist hoch ($W=.81$) und signifikant ($p<.001$). Eine ANOVA für die Daten zeigt eine Abhängigkeit der Mithörschwelle von der Anzahl der Teiltöne der HKTs ($F(3,27)=24.5$,

$p < .001$). Alle signifikanten Unterschiede zwischen den Stimuli sind in Tabelle 7.1 aufgelistet. Die Effektgröße beträgt $\eta^2 = .732$.

Tabelle 7.1.: Auflistung der relevanten, d.h. post-hoc signifikanten (Bonferroni), Unterschiede für die Mithörschwelle

Stimuli	p (Bonferroni)
HKT(1) \leftrightarrow HKT(2)	.022
HKT(1) \leftrightarrow HKT(4)	.002
HKT(1) \leftrightarrow HKT(8)	.001
HKT(2) \leftrightarrow HKT(8)	.021

7.4.3. Diskussion

Die Mithörschwelle für die HKTs wächst mit der Anzahl der Teiltöne, d.h. die Teiltöne sind sukzessive schlechter wahrnehmbar. Die Differenz der Mithörschwelle zwischen HKT(1) und HKT(2) beträgt 1,70 dB, für HKT(2) und HKT(4) 1,05 dB sowie für HKT(4) und HKT(8) 0,4 dB.

Um die Mithörschwelle für die Stimuli bei gleichem Teiltonpegel zu erhalten, muss sie um $10 \log_{10}(k)$ dB reduziert werden, da der Pegel der Stimuli bei Verdopplung der Teiltonanzahl ebenfalls um 3 dB abgesenkt wird. Dieser Fall ist zusätzlich in Abbildung 7.2 eingezeichnet.

Ähnliche Ergebnisse wie die hier gefundenen werden von Schafer und Gales [SG49] präsentiert. Sie untersuchen breitbandig maskierte Tonkomplexe, bei denen alle Teiltöne den gleichen Pegel besitzen und können zeigen, dass die Tonkomplexe auch dann noch wahrnehmbar sind, wenn jeder Teilton allein nicht mehr wahrnehmbar ist. Für Tonkomplexe mit zwei Teiltönen beträgt der Unterschied für die Detektion zwischen Teilton und Tonkomplex etwa 0–2 dB, für Tonkomplexe mit vier bis acht Teiltönen 0–3 dB. Das trifft auch für die hier gemessene Mithörschwelle der HKTs zu.

Green und andere [GML59] zeigen darüber hinaus, dass Tonkomplexe mit 16 Teiltönen, für die die Detektierbarkeit jedes Teiltönen vergleichbar ist, etwa um 6 dB besser wahrgenommen werden können als die (einzelnen) Teiltöne. Beide Untersuchungen führen den „Gewinn“ darauf zurück, dass die Detektion von Tonkomplexen umso besser funktioniert, umso mehr Frequenzgruppen (=auditorische Kanäle) des Zuhörers durch den zu detektierenden Stimulus angeregt werden. Dieser Effekt wird auch als „Informations-Addition“ bezeichnet.

Nach Grose and Hall [GH97] kann die Abhängigkeit der Detektionsschwelle⁶² $L_M(k)$ eines HKTs von der Anzahl der Teiltöne k bei konstantem Pegel der

⁶² Detektionsschwelle und Mithörschwelle werden hier synonym gebraucht.

Teiltöne $L_l(k)$ folgendermaßen angegeben werden:

$$(7.2) \quad L_M(k) = L_{M_t}(k=1) - 10 \log_{10}(\sqrt{k}) \text{ dB} \quad \text{für } L_l(k) = \text{konstant.}$$

Dabei ist $L_{M_t}(k=1)$ die Mithörschwelle für den HKT($k=1$). Sie ist, so wie hier verwendet, eine alternative Bezeichnung des Verdeckungsmaßes $a_v(f)$ für den HKT($k=1$). Damit gilt unter Berücksichtigung von Gleichung A.1 für das Verdeckungsmaß

$$(7.3) \quad L_{M_t}(k=1) = a_v(f) = -2,5 \text{ dB.}$$

Gleichung 7.2 gilt unter der Voraussetzung, dass alle Teiltöne den gleichen Pegel aufweisen. Für die Versuchsstimuli werden jedoch nach Gleichung 7.1 die Teiltonpegel $L_l(k)$ der HKTs bei Verdopplung der Teiltonzahl k um $10 \log_{10}(k)$ dB abgesenkt. Das muss in Gleichung 7.2 berücksichtigt werden, indem dieser Term addiert wird:

$$(7.4) \quad \begin{aligned} L_M(k) &= L_{M_t}(k=1) - 10 \log_{10}(\sqrt{k}) + 10 \log_{10}(k) \\ &= L_{M_t}(k=1) + 10 \log_{10}(\sqrt{k}). \end{aligned}$$

Der Verlauf der Gleichung 7.4 ist in Abbildung 7.2 den gemessenen Daten als „Multiband-Energiemodell“ gegenüber gestellt. Die Bezeichnung ist an dieser Stelle der Diskussion vorweg genommen, eine ausführlichere Erläuterung dazu erfolgt später.

Die gemessenen Mithörschwellen stimmen für HKTs mit $k=1, 2, 4$ sehr gut mit dieser Theorie überein. Lediglich für den HKT mit $k=8$ passt die Vorhersage nicht zu den Messungen. Er wird weniger maskiert (d.h. bei einem geringeren Pegel detektiert) als durch Gleichung 7.4 vorhergesagt.

Für die drei HKTs mit $k \leq 4$ beträgt der Abstand der einzelnen Teiltöne zwei oder mehr Frequenzgruppen. Für den HKT(8) liegen die Teiltöne für die höheren Harmonischen hingegen im Bereich einer Frequenzgruppe. Es kann daher sein, dass die Energie von mehr als einem Teilton innerhalb einer Frequenzgruppe addiert und der Tonkomplex somit weniger maskiert wird als erwartet.

Die vorliegenden Untersuchungen mit den HKTs haben das Ziel, den Beitrag der einzelnen auditorischen Kanäle zur Gesamtwahrnehmung der THK zu untersuchen. Analog zu den hier vorgestellten Detektionsexperimenten existieren zahlreiche weitere, unter anderem z.B. von Grose und Hall [GH97]. Sie zeigen, dass alle Signalkomponenten (Teiltöne) gleich und unabhängig zur Detektionsschwelle beitragen, solange die Teiltöne in einen eigenen auditorischen Kanal fallen und der SNR⁶³ für alle Teiltöne gleich ist.

⁶³ SNR = „Signal to Noise Ratio“, d.h. Verhältnis von tonalem Anteil zum Rauschen.

Für unterschiedliche SNR in den einzelnen auditorischen Kanälen zeigt Marill [Mar56], dass nur der Kanal mit dem besten SNR für die Detektion wichtig ist.

Buus und Kollegen [BSFZ86] können mit Detektionsexperimenten von einzelnen Tönen mit verschiedenen Frequenzen (220 Hz, 110 Hz, 3850 Hz) und komplexen (harmonischen) Tönen mit 18 Harmonischen in gleichmäßig verdeckendem Rauschen zeigen, dass die Detektion dieser Signale am besten durch ein Multiband-Energie-Modell beschrieben werden kann. Das Modell macht die Annahme, dass ein menschlicher Detektor (Hörer) eine ungewichtete Summe über ein optimal ausgewähltes Set von einzelnen auditorischen Kanäle zur Detektion benutzt. Die verschiedenen, einzelnen Töne können allerdings nicht unterschiedlich gewichtet werden. Die Optimierung besteht vielmehr darin, die Kombination der Kanäle bestmöglich an die Situation anzupassen.

Mit diesem Modell können die Daten der meisten Detektionsexperimente angemessen erklärt und vorhergesagt werden. Unter der Annahme, dass die THK auf noch genauer zu bestimmende Weise mit der Mithörschwelle zusammen hängt, muss dieses Modell bei der Betrachtung der THK von HKTs im Weiteren Beachtung finden.

7.5. Variation der Verfahren zur Messung der THK

Um die Einflüsse der experimentellen Parameter auf die Messung der THK (Kapitel 6.2) zu untersuchen, werden die folgenden drei Messdesigns verwendet:

1. *Einzelne Tracks* (ET): Für jeden Teststimulus wird einzeln und nacheinander in einer über die Probanden ausbalancierten Reihenfolge das adaptive THK-Messverfahren durchgeführt.
2. *Interleaved Tracks* (IT): Das adaptive THK-Messverfahren wird interleaved durchgeführt, d.h. die Tracks für alle Teststimuli werden gleichzeitig, ineinander verschachtelt, bewertet.
3. *Blocked Tracks* (BT): Alle Teststimuli werden auch als Vergleichsstimuli verwendet. Da in diesem Fall $4 \cdot 4 = 16$ Tracks bewertet werden müssen, steigt der Messaufwand gegenüber den zuvor genannten Messdesigns quadratisch an. Die 16 Tracks werden interleaved in zwei Messreihen, ausbalanciert über alle Probanden, durchgeführt.

Die Messungen mit den entsprechenden Designs sind in Tabelle 7.2 aufgelistet.

Bei dem *Design ET* können die Probanden dem zu beurteilenden Kriterium in jedem einzelnen Track leicht folgen. Dadurch kann der Effekt auftreten, dass die Probanden ihr Beurteilungskriterium im Verlauf einer Messung den adaptierten

Tabelle 7.2.: Messdesigns für die Messungen der THK. Der HKT($k=1$) ist als „Ton“ bezeichnet, wenn er als THK-Referenzsignal verwendet wird. „Dauer“ bezeichnet die mittlere Messdauer für die entsprechende(n) Messung(en) pro Proband.

Nr.	Design	Startpegel	Teststimulus	Vergleichsstimulus	Dauer
1	ET	$2 \times$ zufällig	HKT($k=1;2;4;8$)	Ton	10 min
2	ET	$2 \times$ zufällig	Ton	HKT($k=1;2;4;8$)	10 min
3	ET	± 10 dB	HKT($k=1;2;4;8$)	Ton	10 min
4	ET	± 10 dB	Ton	HKT($k=1;2;4;8$)	10 min
5	IT	zufällig	HKT($k=1;2;4;8$)	Ton	5 min
6	IT	zufällig	Ton	HKT($k=1;2;4;8$)	5 min
7	IT	± 10 dB	HKT($k=1;2;4;8$)	Ton	10 min
8	IT	± 10 dB	Ton	HKT($k=1;2;4;8$)	10 min
9	BT	zufällig	HKT($k=1;2;4;8$)	HKT($k=1;2;4;8$)	20 min

Stimuli anpassen. Diese Anpassung ist hier nicht gewünscht, da sie vom Verlauf der Messung abhängt und das Ergebnis beeinflusst.

Hinzu kommt, dass der Vergleichsstimulus (Messungen 1 und 3) bzw. der Teststimulus (Messungen 2 und 4) in jedem Track derselbe ist. Durch die Häufung dieses Stimulus kann ebenfalls ein Effekt hervorgerufen werden, der das Ergebnis beeinflusst und daher unerwünscht ist.

Das *Design IT* verhindert, dass die Probanden den einzelnen Tracks ohne weiteres folgen können. Dadurch wird die systematische Anpassung der Probandenurteile im Verlauf eines Tracks minimiert. Allerdings besteht auch für dieses Messdesign das Problem, dass der Vergleichsstimulus (Messungen 5 und 7) bzw. der Teststimulus (Messungen 6 und 8) deutlich öfter präsentiert wird als der Teststimulus bzw. der Vergleichsstimulus.

Diese Problematik wird durch das *Design BT* verhindert: alle Stimuli werden gleich häufig präsentiert. Darüber hinaus weiß der Proband nicht, welcher der beiden Stimuli bei der aktuellen Bewertung der Vergleichsstimulus bzw. der Teststimulus ist. Auf diese Weise ist es für die Probanden unmöglich, dem Verlauf der Bewertungen zu folgen. Stattdessen muss der Proband für jede neue Bewertung ausschließlich anhand seines internen Vergleichskriteriums urteilen, da keine zusätzlichen Hinweise, die sich aus dem Messdesign ergeben, zur Verfügung stehen. Somit werden die diskutierten systematischen Fehler („Bias“) bei den Bewertungen bestmöglich vermieden.

Für die Startpegel werden verschiedene Ansätze verfolgt: Messungen 3, 4, 7 und 8 werden zwei mal mit verschiedenem Startpegel (± 10 dB zur nominellen, d.h. berechneten THK für den HKT(1)) durchgeführt. Das entspricht dem Vorgehen in den bisherigen Untersuchungen.

Messungen 1 und 2 werden ebenfalls zwei mal durchgeführt, der Startpegel wird aber zufällig aus dem Intervall $[-10 \text{ dB}; +10 \text{ dB}]$ relativ zur nominellen THK gewählt. Der Startpegel und somit das Ergebnis der Messung (PSE der THK) sind in diesem Fall beliebige reelle Zahlen.

Eine Vereinfachung dieses Ansatzes sind Messungen 5 und 6. Hier werden die Messungen nur einmal durchgeführt und nicht wiederholt. Auf diese Weise wird auch der Startpegel für die Messung 9 mit dem Messdesign BT festgelegt.

Die Reihenfolge der Messungen 1 bis 8 aus Tabelle 7.2 erfolgt (soweit möglich) ausbalanciert über die Probanden.

7.6. Ergebnisse des Verfahrensvergleich

Prinzipiell können die Messungen nach Tabelle 7.2 in drei verschiedene Klassen eingeteilt werden, die im Folgenden getrennt ausgewertet werden:

1. Die zu untersuchenden Signale werden als Teststimuli eingesetzt. Sie werden anhand des THK-Referenzsignals (als Vergleichsstimulus) bewertet. Das gilt für die Messungen 1, 3, 5 und 7 und entspricht dem in Kapitel 6.2 vorgestellten und bisher verwendeten Verfahren.
2. Das THK-Referenzsignal wird als Teststimulus, die zu untersuchenden Signale als Vergleichsstimuli verwendet. In diesem Fall werden die zu untersuchenden Signale adaptiv verändert und auf die gleiche THK wie das konstante Referenz-Signal eingestellt. Messungen 2, 4, 6 und 8 werden auf diese Weise durchgeführt.
3. Die Messungen mit Blocked Tracks vereinen die beiden zuvor genannten Klassen. Sie werden jedoch ebenfalls getrennt ausgewertet, da ihre Komplexität erheblich über den beiden anderen Klassen liegt.

Im Kapitel 6 werden die Daten zunächst konservativ mit parameterfreien Methoden statistisch untersucht, gefolgt von einer Auswertung auf Intervallskalenniveau. Im Anschluss daran werden die Ergebnisse gegenübergestellt. Die Interpretation der Messdaten für die beiden Skalenniveaus stehen in Einklang und ergänzen sich. In keinem Fall sind gegenläufige Tendenzen zu erkennen⁶⁴.

Aus diesem Grund wird die Datenanalyse im Folgenden ausschließlich auf Intervallskalenniveau durchgeführt. Die Aussagekraft der statistischen Verfahren ist größer und es stehen mehr Analysemöglichkeiten zur Verfügung. Insbesondere kann problemlos eine ANOVA durchgeführt werden, die schnell die wichtigsten Zusammenhänge der Messdaten aufdecken kann.

⁶⁴ Siehe dazu auch die Diskussion im Kapitel 6 und die Ergebnisse im Anhang F.3.

Um die Vergleichbarkeit der im Folgenden präsentierten Messungen mit den zuvor durchgeführten zu gewährleisten, wird als Konkordanzmaß wieder Kendall's- W verwendet.

Die Messungen 1–4 sowie 7 und 8 werden zwei mal mit verschiedenen Startpegeln wiederholt (Test und Retest). Für die weitere Auswertung sind Test und Retest durch arithmetische Mittelung zusammen gefasst, da analog zu Kapitel 6.4 kein nachweisbarer Effekt des Startpegels vermutet wird. Eine Ausführlichere Diskussion erfolgt, nachdem die entsprechenden Messungen ausgewertet worden sind.

7.6.1. Bewertung der HKTs als Teststimuli

Zunächst werden die Messungen präsentiert, bei denen die THK als Teststimulus bewertet wird, d.h. die Messungen 1, 3, 5 und 7 aus Tabelle 7.2. Diese vier Messungen sind im Folgenden verkürzt als „Messung Teststimulus“ bezeichnet.

Das adaptive Verfahren liefert die Pegeldifferenz L_D zwischen der maßgeblichen Differenz für den Vergleichsstimulus am PSE ΔL_{RefPSE} und der maßgeblichen Differenz des Teststimulus ΔL_T , vgl. Kapitel 6.3.3.

Für die vorliegenden Messungen ist der Vergleichsstimulus in allen Fällen das THK Referenzsignal. Im Folgenden wird nur der Ton des Referenzsignals betrachtet, da das Rauschen für alle Stimuli identisch ist. Dann gilt Gleichung 6.11. Unter Verwendung von Gleichung 6.12 ergibt sich so $\Delta L_{RefPSE} = L_D + 16,3 \text{ dB}$.

Die gemessenen THK-Daten sind auf diese Weise in Abbildung 7.3 aufgetragen. Die gestrichelte Gerade bei 16,3 dB symbolisiert die THK des Vergleichsstimulus, d.h. $L_D=0 \text{ dB}$. Die Konkordanz der Urteiler W ist für die vier Messungen vergleichbar groß $.57 \leq W \leq .65$.

Der Verlauf der Daten ist für alle Messungen gleich: die THK steigt mit zunehmender Anzahl von Teiltönen deutlich an, das Ausmaß des Anstiegs ist für die Messdesigns jedoch unterschiedlich.

Eine zweifaktorielle ANOVA (messwiederholt) über die vier Messungen mit den beiden Faktoren „Anzahl der Teiltöne“ k und „Messdesign“ zeigt einen Effekt sowohl für die Teiltöne⁶⁵ ($F(1,165,30)=9.4$, $p=.008$, $\eta^2=.486$) als auch für das Messdesign ($F(3,30)=7.7$, $p=.001$, $\eta^2=.435$).

Post-hoc durchgeführte, gepaarte T -Tests (Bonferroni) offenbaren signifikante Unterschiede für das Messdesign (Tabelle 7.3). Die THK für Messung 1 wird signifikant kleiner eingeschätzt als für die anderen drei Messungen (mittlere THK für die vier Messungen: 1,98 dB; 2,67 dB; 3,23 dB; 3,23 dB). Dabei muss

⁶⁵ Weil der Mauchly-Test eine Verletzung der Sphärizitätsannahme anzeigt (d.h. die Varianzen der Differenzen zwischen jeweils zwei Faktorstufen sind nicht homogen), muss eine Korrektur der Freiheitsgrade vorgenommen werden. Diese wird hier und im Folgenden immer nach dem Verfahren von „Greenhouse-Geisser“ durchgeführt [SPS02].

7. Verfahrensvergleich zur Bestimmung der Tonhaltigkeit

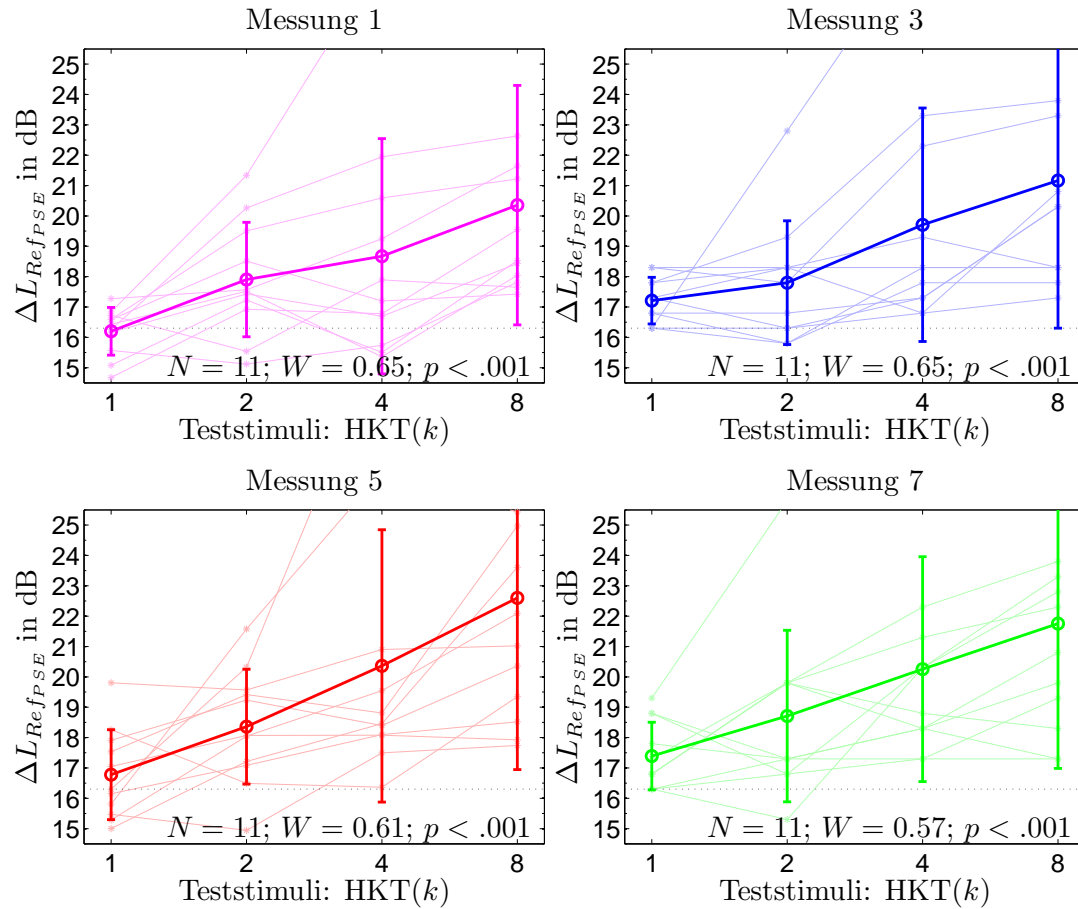


Abbildung 7.3.: Tonhaltigkeit der HKTs für Messungen 1, 3, 5 und 7 für die Teststimuli, skaliert anhand des THK-Referenzsignals als Vergleichsstimulus (HKT(1)). Die dünnen Linien geben den individuellen Verlauf der THK für die elf Probanden wieder, die dicken Linien repräsentieren die Mittelwerte samt Standardabweichung. Ein Proband bewertet, verglichen mit den zehn anderen Probanden derart extrem, dass die Daten außerhalb des dargestellten Bereiches liegen.

allerdings berücksichtigt werden, dass für Messung 1 die THK für den HKT(1) deutlich näher am erwarteten Wert von 16,3 dB liegt als für die anderen drei Messungen.

Auch für die Stimuli existieren signifikante Unterschiede (Tabelle 7.4). Der HKT(8) ruft eine größere THK hervor als alle anderen Stimuli.

Der Mittelwert über die vier diskutierten Messungen 1,3,5,7 ist in Abbildung 7.6 eingetragen, um ihn im Folgenden mit den anderen Messungen bzw. Messdesigns vergleichen zu können.

Tabelle 7.3.: Signifikante Unterschiede zwischen den Messdesigns (post-hoc, Bonferroni) bei der Bewertung der HKTs als Teststimuli

Messung	p (Bonferroni)
1 \leftrightarrow 3	.020
1 \leftrightarrow 5	.049
1 \leftrightarrow 7	.005

Tabelle 7.4.: Auflistung der relevanten, d.h. post-hoc signifikanten (Bonferroni), Unterschiede für die HKTs bei der Bewertung als Teststimuli

Stimuli	p (Bonferroni)
HKT(1) \leftrightarrow HKT(8)	.043
HKT(1) \leftrightarrow HKT(8)	.031
HKT(1) \leftrightarrow HKT(8)	.007

7.6.2. Bewertung der HKTs als Vergleichsstimuli

In den Messungen 2, 4, 6, 8 wird das THK-Referenzsignal als Teststimulus verwendet und mittels der zu untersuchenden Signale (als Vergleichsstimulus) skaliert. Diese vier Messungen sind im Folgenden verkürzt als „Messung Vergleichsstimulus“ bezeichnet.

In diesem Fall sind die Verdeckungsmaße für die vier Vergleichsstimuli verschieden⁶⁶, so dass die Messungen nicht unmittelbar miteinander verglichen werden können. Die gemessenen Pegeldifferenzen $L_D(k)$ (siehe Abbildung G.1 im Anhang G) für die drei Vergleichsstimuli HKT(2), HKT(4) und HKT(8) müssen zuvor auf die Pegeldifferenz $L_D(1)$ normiert werden, um diese Abhängigkeit aufzulösen.

Das wird umgesetzt, indem die Verdeckungsmaße $a_v(k)$ nach Gleichung 6.10 berücksichtigt werden. Da hier Vergleichs- und Teststimulus gegenüber der Definition in Kapitel 6.3.3 vertauscht sind, muss das Vorzeichen $L_D(k)$ invertiert werden:

$$(7.5) \quad -L_{D_{kor}}(k) = (L_{Ref_{PSE}}(k) - a_v(k)) - (L_{T_{est}} - a_v(1)).$$

Die Messung der Verdeckungsmaße ist im Kapitel 7.4 beschrieben, die Werte für $a_v(k)$ können direkt aus Abbildung 7.2 entnommen werden. Die so korrigierten Messwerte zeigt Abbildung G.2 im Anhang G.

⁶⁶ Ein ähnliches Problem wird bereits im Kapitel 6.4.9 behandelt - dort sollen Daten für unterschiedliche Vergleichsstimuli (700 Hz und 1850 Hz) in einem Diagramm dargestellt werden.

7. Verfahrensvergleich zur Bestimmung der Tonhaltigkeit

Gleichung 7.5 wird nach $L_{Ref_{PSE}}(k)$ freigestellt. Durch Einsetzen der entsprechenden Werte ergibt sich analog zu Gleichung 6.13

$$(7.6) \quad L_{Ref_{PSE}}(k) = -L_{D_{kor}}(k) + 16,3 \text{ dB.}$$

Die gemessenen THK-Daten sind mithilfe dieser Transformation umgerechnet und in Abbildung 7.4 aufgetragen. Die Gerade bei 16,3 dB symbolisiert die THK des Vergleichsstimulus, d.h. $L_D = L_{D_{kor}} = 0$ dB. Die eingezeichneten Standardabweichungen der gemessenen Daten werden durch Addition der Varianzen für die gemessene THK (L_D) und dem Verdeckungsmaß gebildet (Gauß'sche Fehlerfortpflanzung).

Prinzipiell ergibt sich der gleiche Verlauf der THK wie in den Messungen 1, 3, 5, 7, allerdings sind HKT(2) und HKT(4) in diesem Fall etwas größer. Die Konkordanz der Urteiler ist ebenfalls etwas größer als in der vorhergehenden Messung ($.65 \leq W \leq .81$).

Eine zweifaktorielle ANOVA (messwiederholt) über die vier Messungen zeigt einen deutlichen Effekt für die Anzahl der Teiltöne ($F(3,30)=45.3$, $p < .001$, $\eta^2 = .819$) und einen kleineren Effekt für das Messdesign ($F(3,30)=3.9$, $p = .018$, $\eta^2 = .282$).

Messung 2 zeigt die kleinste mittlere THK, gefolgt von Messungen 4, 8 und 6 (1,98 dB; 2,79 dB; 3,54 dB; 3,57 dB). Diese Unterschiede zwischen den Messungen sind allerdings nicht signifikant auf dem 5%-Signifikanzniveau (post-hoc durchgeführte, gepaarte T -Tests (Bonferroni)).

Die signifikanten Unterschiede zwischen den Teiltönen HKT(k) sind in Tabelle 7.5 aufgelistet.

Tabelle 7.5.: Signifikante Unterschiede zwischen den Stimuli (post-hoc, Bonferroni) für die Bewertung der THK als Vergleichsstimuli

Stimuli	p (Bonferroni)
HKT(1) ↔ HKT(2)	.002
HKT(1) ↔ HKT(4)	<.001
HKT(1) ↔ HKT(8)	<.001
HKT(2) ↔ HKT(4)	.003
HKT(2) ↔ HKT(8)	.003

Der Mittelwert über die vier diskutierten Messungen 2, 4, 6, 8 aus Abbildung 7.4 samt Standardfehler ist in Abbildung 7.6 eingetragen, um ihn mit den folgenden Messungen vergleichen zu können.

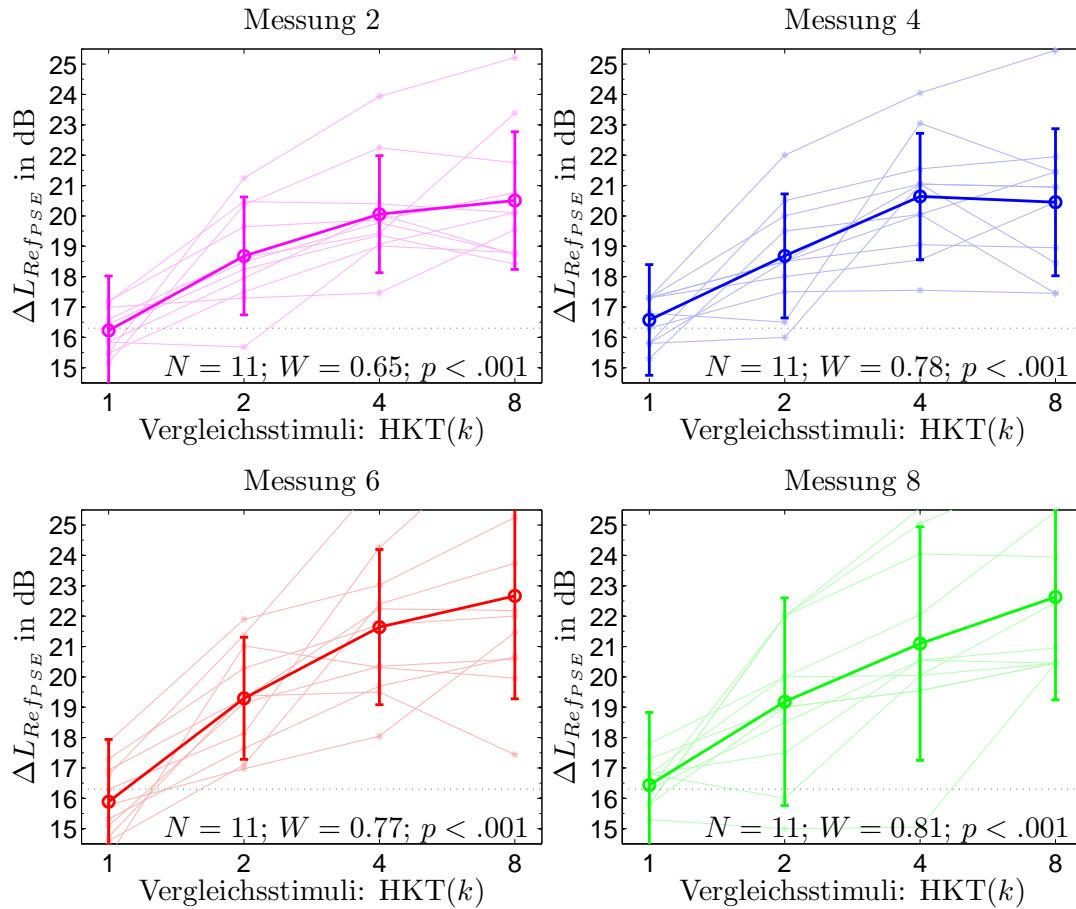


Abbildung 7.4.: Tonhaltigkeit der HKTs für Messungen 2, 4, 6 und 8. Für diese Messungen wird das THK-Referenzsignal als Teststimulus verwendet und anhand der zu untersuchenden Signale (als Vergleichsstimuli) skaliert. Aufgetragen ist die THK über die Vergleichsstimuli, die jeweils um ihr individuelles Verdeckungsmaß $a_v(k)$ korrigiert sind.

Die dünnen Linien geben den individuellen Verlauf für die elf Probanden wieder, die dicken Linien repräsentieren die Mittelwerte samt Standardabweichung. Die angegebene, zusammengefasste Standardabweichung berücksichtigt sowohl die Standardabweichung für die Messung der Schwelle als auch der THK (Gauß'sche Fehlerfortpflanzung).

7.6.3. Vergleich der Bewertung als Vergleichs- bzw. Teststimuli

Die Messungen 1 und 2, 3 und 4, 5 und 6, sowie 7 und 8 sind prinzipiell die gleichen Messungen⁶⁷, die sich lediglich durch die „Richtung“ der Skalierung unterscheiden - die Differenz L_D nach Kapitel 6.3.3 bzw. 7.6.2 hat ein unterschiedliches Vorzeichen.

Da nicht zwangsläufig zu erwarten ist, dass die aufgezählten Messungen die gleichen Ergebnisse liefern, sind diese in Abbildung 7.5 grafisch gegenübergestellt.

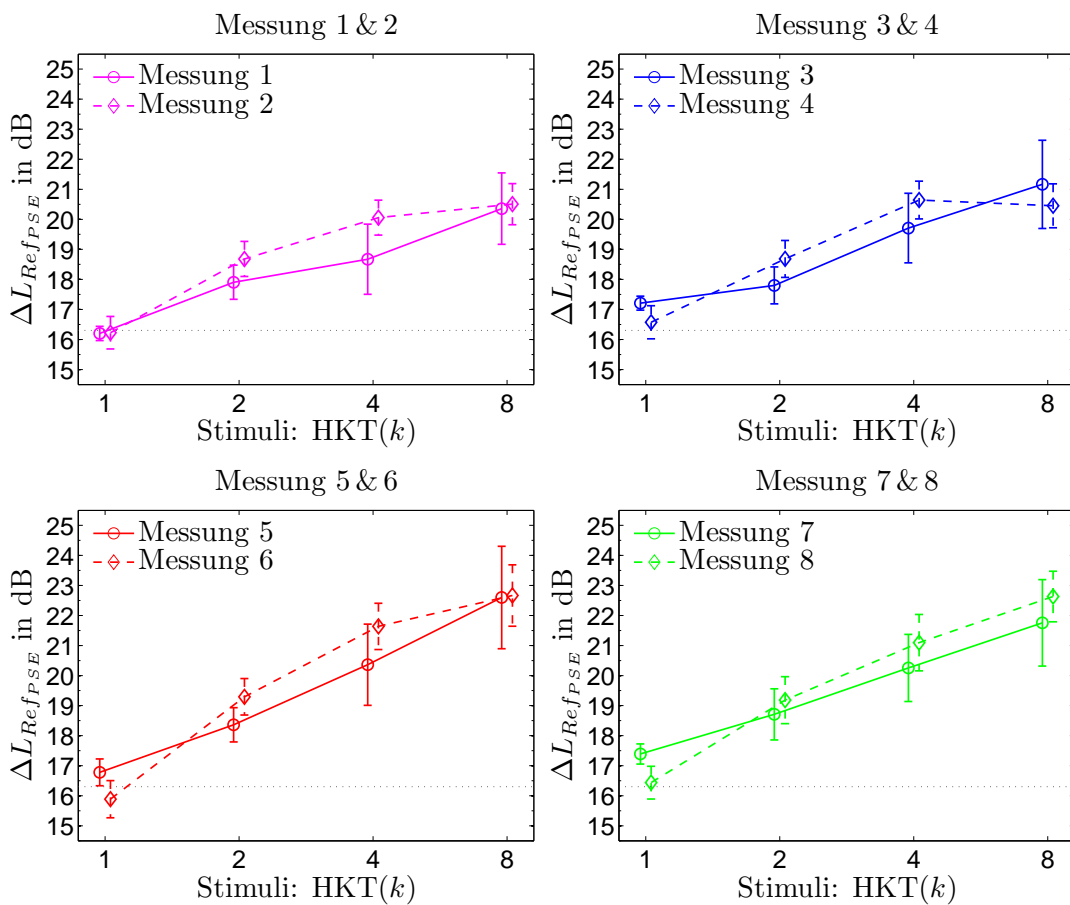


Abbildung 7.5.: Gegenüberstellung der zu vergleichenden Messungen (1 und 2, 3 und 4, 5 und 6, sowie 7 und 8). Aufgetragen sind Mittelwerte samt Standardfehler für die THK.

Die Gegenüberstellung zeigt ansatzweise eine Art Hysterese für die Messung: Die relative Skalierung der HTKs anhand des HKT(1) liefert kleinere Werte

⁶⁷ im Folgenden als „identische Messungen“ bezeichnet

als für die Skalierung anhand des HKT(8). Eine zweifaktorielle ANOVA (messwiederholt) kann diesen Effekt auf dem 5% Signifikanzniveau jedoch für keine der vier Diagramme in Abbildung 7.5 nachweisen, d.h. die gegenüber gestellten Messungen können als gleich betrachtet werden.

Die Standardabweichungen (bzw. die Standardfehler in Abbildung 7.5) sind größer für die Messungen mit gerader Ordnungszahl (3,0 dB) als für die Messungen mit ungerader Ordnungszahl (1,9 dB).

Für die Messungen mit ungerader Ordnungszahl nehmen sie mit der Anzahl der Teiltöne der HKTs zu. Das ist damit zu erklären, dass der Abstand zwischen den Teststimuli und den Vergleichsstimuli größer wird und die Unsicherheit der Probanden bei der Bewertung dadurch wächst.

Für die Messungen mit geraden Ordnungszahlen ist dieser Effekt ebenso vorhanden, allerdings deutlich kleiner. Zusätzlich wird die Unsicherheit bei der Messung der THK durch die Unsicherheit bei der Messung der Mithörschwelle überlagert. Aus diesem Grund sind die Standardabweichungen für HKT(1) der Messungen mit geraden Ordnungszahlen größer als für die Messungen mit ungeraden Ordnungszahlen, für HKT(2) etwa vergleichbar und für HKT(4) und HKT(8) geringer.

7.6.4. Vergleich der Messdesigns ET und IT sowie der Startpegel

Der Einfluss der Startpegel auf die Messergebnisse wird geprüft, indem jeweils die Messungen mit festem Startpegel (3, 4, 7, 8) den Messungen mit zufälligen Startpegeln (1, 2, 5, 6) gegenübergestellt werden, vgl. Tabelle 7.2. Eine dreifaktorielle ANOVA zeigt für diesen Zusammenhang keinen signifikanten Effekt auf dem 5%-Signifikanzniveau ($F(2,10)=2.7$, $p=.13$, $\eta^2=.212$), d.h. die Startpegel haben keinen systematischen Effekt auf die Messungen.

Um zusätzlich zu analysieren, ob es einen Unterschied zwischen den Messungen mit nur einer Messwiederholung (5 und 6) und den analogen Messungen mit zwei Messwiederholungen (7 und 8) gibt, wird jeweils eine ANOVA für die Messungen 5 und 7, sowie 6 und 8 durchgeführt. Es zeigt sich in beiden Fällen kein Effekt auf dem Signifikanzniveau von 5%.

Die Messdesigns werden verglichen, indem die Messungen entsprechend gruppiert werden: ET (Messungen 1–4) bzw. IT (Messungen 5–8). Dafür werden die Ergebnisse der beiden Messdesigns zunächst in Abbildung 7.6 visualisiert und den beiden bereits diskutierten Kurven gegenübergestellt.

Die zu Grunde liegenden Daten für die Messdesigns sind identisch zu den in den voran gegangenen Abschnitten. Daher sind keine neuen Erkenntnisse bzgl. der Anzahl der Teiltöne zu erwarten. Es ist hingegen interessant zu untersuchen, ob die beiden Messdesigns sich voneinander unterscheiden (d.h. Bias-behaftet

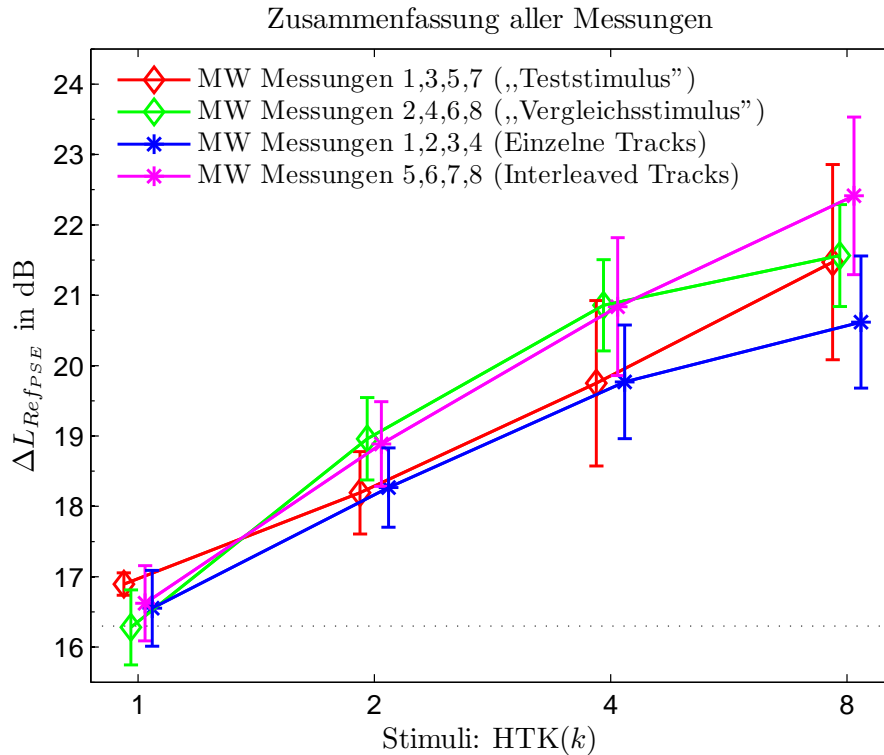


Abbildung 7.6.: Übersicht der Ergebnisse der Messungen 1–8 aus Tabelle 7.2, zusammengefasst anhand der diskutierten Gruppen bzw. Messdesigns. Aufgetragen sind Mittelwerte (MW) und Standardfehler. Die Messungen sind zur besseren Darstellung horizontal gegeneinander verschoben.

sind) oder nicht. Aus diesem Grund wird eine dreifaktorielle ANOVA (messwiederholt) über die Faktoren „Messdesign“ (ET/IT), „identische Messungen“ und „Anzahl der Teiltöne“ durchgeführt.

Es zeigt sich ein signifikanter Effekt des Messdesigns ($F(1,10)=19.5$, $p=.001$, $\eta^2=.661$), d.h. die Mittelwerte für ET (2,50 dB) und IT (3,39 dB) unterscheiden sich überzufällig. Die Wahl des Messdesigns beeinflusst somit die Resultate der THK-Messung.

Den Erwartungen entsprechend, ist der gemittelte Wert der einzelnen Standardfehler als Maß der Unsicherheit der Probanden bei IT (0,85 dB) größer als bei ET (0,66 dB).

Die ANOVA zeigt einen Einfluss der Teiltonanzahl ($F(1.18,11.80)^{68}=22.35$, $\eta^2=.691$), d.h. der Stimuli HKT(k), jedoch keinen Effekt der „identischen Messungen“ ($F(1.24,12.43)^{68}<1$ $\eta^2=.055$).

⁶⁸ Da der Mauchly-Test auf Sphärizität signifikant ist, wird die Greenhouse-Geisser Korrektur angewendet.

7.6.5. Zusammenfassung und Diskussion der Messungen 1–8

Die Ergebnisse der in Tabelle 7.2 aufgeführten Messungen können wie folgt zusammen gefasst werden:

- Die Messdesigns IT und ET unterscheiden sich statistisch belastbar voneinander. Messungen, die mit IT durchgeführt werden, weisen eine größere THK auf. Das Messdesign beeinflusst die gemessene THK somit systematisch und verursacht einen Bias der Messung.
Mit den vorliegenden Daten kann nicht beurteilt werden, welches der beiden Verfahren die THK „besser“ beurteilt. Daher wird die THK im nächsten Abschnitt zusätzlich mit dem „Blocked Design“ ermittelt und mit den bisher erzielten Ergebnissen verglichen.
- Die mittlere Standardabweichung als Maß für die Unsicherheit der Probanden ist für das Messdesign IT größer als für ET. Das ist im Vorfeld der Messungen bereits vermutet worden und bestätigt die Annahme, dass die Probanden durch Verwendung des Messdesigns IT stärker „verwirrt“ werden sollen.
- Im Rahmen der vorgestellten Resultate ist es unerheblich, ob die zu untersuchenden $HKT(k)$ als Test- oder Vergleichsstimulus bewertet werden. Tendenziell existieren kleinere Unterschiede („Hysterese“), die sich jedoch statistisch nicht nachweisen lassen.
Die mittlere Standardabweichung für die Messungen „Vergleichsstimulus“ beträgt $\approx 65\%$ der Messungen „Teststimulus“. Sie ist damit deutlich kleiner.
- Für den hier vorgenommenen Verfahrensvergleich ist es von Interesse, die Bewertung der THK sowohl als Test-, als auch als Vergleichsstimulus durchzuführen, um die Unterschiede zu quantifizieren. Für die praktische Anwendung ist es hingegen i.d.R. nicht möglich, das zu bewertende Signal als Vergleichsstimulus zu verwenden. In diesem Fall wird das zu bewertende Signal fast immer als Teststimulus skaliert.
- Die Wahl des Startpegels hat keinen Einfluss auf die Messung der THK. Insbesondere kann kein Unterschied ermittelt werden, ob die Messungen mit zwei unterschiedlichen, festen Startpegeln durchgeführt und gemittelt werden, oder nur eine einzige Messung mit zufälligem Startpegel ausgeführt wird.
Die Messzeit, und damit der Aufwand für die Probanden, kann somit durch Verwendung einer einzelnen Messung mit zufälligem Startpegel gegenüber allen anderen Messungen aus Tabelle 7.2 halbiert werden.

7. Verfahrensvergleich zur Bestimmung der Tonhaltigkeit

Die Stimuli, d.h. die $HKT(k)$, unterscheiden sich bei gemeinsamer Betrachtung über alle acht Messungen hinweg signifikant, siehe Tabelle 7.6.

Tabelle 7.6.: Signifikante Unterschiede zwischen den Stimuli über alle Messungen (post-hoc, Bonferroni).

Stimuli	p (Bonferroni)
$HKT(1) \leftrightarrow HKT(2)$.005
$HKT(1) \leftrightarrow HKT(4)$.006
$HKT(1) \leftrightarrow HKT(8)$.003
$HKT(2) \leftrightarrow HKT(4)$.044
$HKT(2) \leftrightarrow HKT(8)$.005
$HKT(4) \leftrightarrow HKT(8)$.002

Die Mittelwerte samt Standardfehler der THK für die $HKT(k)$ enthält Tabelle 7.7. Sie sind ein wesentliches Ergebnis der voran gegangenen Messungen.

Tabelle 7.7.: Mittelwerte und Standardfehler der HKTs über alle acht Messungen. Sämtliche Angaben in dB.

Stimulus	Mittelwert	Standardfehler
HKT(1)	0,29	0,11
HKT(2)	2,28	0,47
HKT(4)	4,00	0,84
HKT(8)	5,22	0,99

Der Mittelwert von 0,29 dB für den HKT(1) ermöglicht eine Abschätzung der Genauigkeit des Messverfahrens. Bei idealem Ausgang der Messung sollte dieser Wert 0 dB betragen. Es liegt somit eine Abweichung unter 0,3 dB vor.

Alle adaptiven Messungen werden mit einer minimalen Schrittweite von 2 dB durchgeführt. Zudem sind die Startpegel bei der Hälfte der Messungen (Nr. 3, 4, 7, 8) feststehend. Daraus ergibt sich, dass es nur eine begrenzte Anzahl von diskreten Messergebnissen für die einzelnen Probanden geben kann. Der Unterschied der Ergebnisse muss immer ein ganzzahliges Vielfaches der minimalen Schrittweite sein.

Erst durch die Mittelung über die Probanden hinweg wird ein genaueres Ergebnis, welches zwischen den diskreten, einzelnen Messwerten liegt, möglich. Somit ist die Abweichung von 0,29 dB zwischen dem gemessenen und dem erwarteten Wert der THK für den HKT(1) in einem vertretbaren Bereich.

7.6.6. Bewertung der HKTs mit Blocked Tracks

Der vorige Abschnitt zeigt, dass sich einige der durchgeführten Messungen signifikant unterscheiden, obwohl die gleichen Stimuli skaliert werden. Um zu untersuchen, ob bzw. welche Messungen mit systematischen Fehlern behaftet sind, wird im folgenden die Bewertung mit Blocked Tracks (BT) wiederholt.

Die Rohdaten der 16 einzelnen Messungen sind im Anhang, Tabelle G.1 und Abbildung G.3 dargestellt. Die Daten können wahlweise über die Vergleichsstimuli oder die Teststimuli aufgetragen werden. Da die zweite Darstellung leichter zu interpretieren ist, wird diese verwendet.

An dieser Stelle tritt das gleiche Problem auf wie in Kapitel 7.6.2. Das Verdeckungsmaß ist für die vier $HKT(k)$ jeweils unterschiedlich. Daher können die Messwerte nicht direkt miteinander verglichen, sondern müssen zuvor geeignet transformiert werden. Dazu werden die Verdeckungsmaße der einzelnen $HKT(k)$ berücksichtigt, indem jeweils alle $HKT(k > 1)$ auf den $HKT(1)$ normiert werden.

Die einzelnen Schritte dieser Transformation sind in Anhang G.2 aufgeführt, das gemittelte Endergebnis zeigt Tabelle 7.8 und Abbildung 7.7.

Tabelle 7.8.: Gemessene THK mit Blocked Tracks. Dargestellt sind die Mittelwerte samt Standardfehler über die vier Vergleichsstimuli für elf Probanden. Alle Angaben in dB.

Teststimulus	THK (L_D)	THK (ΔL_{RefPSE})
HKT(1)	0,66 ± 0,84	16,96 ± 0,84
HKT(2)	1,74 ± 0,86	18,04 ± 0,86
HKT(4)	2,30 ± 0,94	18,60 ± 0,94
HKT(8)	3,41 ± 1,11	19,71 ± 1,11

Der Verlauf der so ermittelten THK entspricht dem Verlauf der Messung aus dem vorigen Abschnitt (Messungen 1–8). Die absoluten Werte sind für diese Messung jedoch bis zu 2 dB kleiner, während der Standardfehler leicht größer ist.

Beide Punkte zusammen genommen bestätigen die eingangs des Kapitels aufgestellten Annahmen. Die „Verwirrung“ der Probanden wird durch das Messverfahren BT erhöht. Dadurch werden sie einerseits unsicherer in ihren Urteilen, sodass der Standardfehler für die Mittelwerte steigt. Andererseits kann die Unsicherheit die Probanden dazu verleiten, mehr Urteile zufällig zu fällen. Das wiederum führt dazu, dass L_D als Maß für die gemessene THK um den Nullwert schwankt, wodurch sich ein kleinerer Mittelwert und damit ein geringerer Effekt für die THK zeigt.

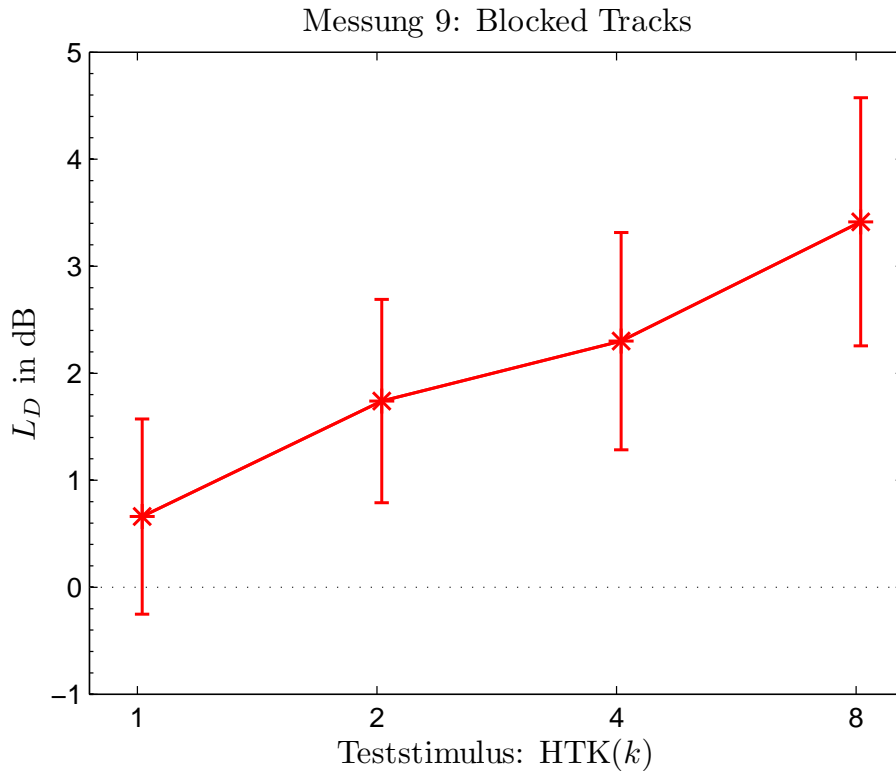


Abbildung 7.7.: Transformierte THK-Daten für Blocked Tracks. Die Daten der Abbildung G.3 werden mittels der individuellen Verdeckungsmaße der HKT(k) auf den HKT(1) normiert und über die Vergleichsstimuli gemittelt. Aufgetragen sind diese Mittelwerte über die elf Probanden samt Standardfehler. Der Standardfehler berücksichtigt sowohl den Anteil der Messung der Mithörschwelle als auch den Anteil der Messung der THK (Gauß'sche Fehlerfortpflanzung).

7.7. Modellierung der Tonhaltigkeit für harmonisch komplexe Töne

Die Ergebnisse der drei zuvor durchgeführten Untersuchungen zur THK von HKTS sollen im Folgenden zusammen geführt und vereinheitlicht werden.

Dazu ist die gemessene THK THK_{gem} in den Tabellen 7.9, 7.10 und 7.11, getrennt für die drei Untersuchungen, erneut dargestellt.

In der ersten Spalte aller Tabellen steht die Anzahl der Teiltöne, hier als n bezeichnet, für die HKT(n).

Die gemessene THK_{gem} für die HKT(1) ist in allen Tabellen in Klammern gesetzt, da sie idealerweise 0 dB betragen sollte und für die weiteren Analysen

Tabelle 7.9.: Gemessene THK für die HKTs „ohne Energieausgleich“ (Abbildung 7.8). Alle Angaben in dB.

n	THK_{gem}	ΔL_l	$a_{i_{theo}}$	$a_{i_{gem}}$	X_{theo}	X_{gem}
1	(0,3)	0	0	0	0	0
2	5,1	0	-1,5	-1,30	6,6	6,4
4	8,2	0	-3,0	-3,25	11,2	11,45
8	12,4	0	-4,5	-5,85	16,9	18,25
gewichteter Mittelwert b_N :					5,94	6,07

Tabelle 7.10.: Mit Blocked Tracks gemessene THK für die HKTs (Abbildung 7.9). Alle Angaben in dB.

n	THK_{gem}	ΔL_l	$a_{i_{theo}}$	$a_{i_{gem}}$	X_{theo}	X_{gem}
1	(0,7)	0	0	0	0	0
2	1,7	-3	-1,5	-1,30	6,2	6,0
4	2,3	-6	-3,0	-3,25	11,3	11,55
8	3,4	-9	-4,5	-5,85	16,9	18,25
gewichteter Mittelwert b_N :					5,82	5,95

Tabelle 7.11.: Gemessene THK (Messungen 1–8 der Tabelle 7.2) für die HKTs (Abbildung 7.10). Alle Angaben in dB.

n	THK_{gem}	ΔL_l	$a_{i_{theo}}$	$a_{i_{gem}}$	X_{theo}	X_{gem}
1	(0,3)	0	0	0	0	0
2	2,3	-3	-1,5	-1,30	6,8	6,6
4	4,0	-6	-3,0	-3,25	13,0	13,25
8	5,3	-9	-4,5	-5,85	18,8	20,15
gewichteter Mittelwert b_N :					6,52	6,65

nicht benötigt wird.

Die dritte Spalte der Tabellen enthält die maßgebliche Differenz ΔL_l der *einzelnen Teiltöne* der HKTs. Pro HKT(n) ist die maßgebliche Differenz für alle Teiltöne per Definition identisch, zwischen den HKT(n) allerdings für die Messungen dieses Kapitels mit $10 \log_{10}(n)$ dB abfallend. Die maßgebliche Differenz der Teiltöne wird daher jeweils auf die maßgebliche Differenz des HKT(1) bezogen.

In der vierten und fünften Spalte ist die Änderung der Mithörschwelle a_i beim Vorhandensein mehrerer Teiltöne aufgrund des Multiband-Energiemodells („Informations-Addition“) eingetragen. Die Mithörschwelle a_v der einzelnen Teiltöne

ist in der maßgeblichen Differenz der Teiltöne ΔL_l enthalten. Sie ist per Definition konstant und wird hier daher nicht dargestellt.

Die Informations-Addition wird im Folgenden relativ zum HKT(1) angegeben (d.h. $a_i(1)=0$ dB). Dabei wird zwischen der aufgrund des Multiband-Energiemodells vorhergesagten Informations-Addition $a_{i_{theo}} = -10 \log_{10}(\sqrt{n})$ und der gemessenen Informations-Addition $a_{i_{gem}}$ unterschieden.

Die gemessene Informations-Addition wird aus der experimentell ermittelten Änderung der Mithörschwelle für die HKT(n) bestimmt, vgl. Abbildung 7.2 „MH- $10 \log_{10}(\sqrt{n})$ dB“.

Im Folgenden wird die Differenz X der gemessenen THK zur Mithörschwelle eines einzelnen Teiltones berechnet als

$$(7.7) \quad X = THK_{gem} - \Delta L_l - a_i.$$

Diese Differenz ist in den Tabellen für die anhand des Multiband-Energiemodells vorhergesagte Informations-Addition $a_{i_{theo}}$ als X_{theo} und für die gemessene Informations-Addition als X_{gem} bezeichnet. Die Differenz X ist anschaulich der Pegel über der Mithörschwelle für einen Teilton.

Da die Teiltöne nach Voraussetzung alle die gleiche Mithörschwelle besitzen, braucht der konkrete Teilton nicht näher spezifiziert zu werden. Die Anzahl der Teiltöne muss hingegen berücksichtigt werden. Eine angemessene Beschreibung ergibt sich durch Verwendung des folgenden Ausdrucks

$$(7.8) \quad X = a_N \log_2(n).$$

Der Term a_N ist eine Konstante, die es zu bestimmen gilt. Für die vorliegenden Messungen liegt er zwischen 5,6 dB und 6,7 dB.

Für die Messung der HKTs „ohne Energieausgleich“ (Tabelle 7.9) ergibt sich a_N als gewichteter Mittelwert für die theoretisch vorhergesagte Mithörschwelle zu 5,9 dB und für die gemessene Mithörschwelle zu 6,1 dB. Aus Tabelle 7.10 ergibt sich a_N für die Blocked Tracks zu 5,8 dB (theoretisch) bzw. 6,0 dB (gemessen). a_N für beide Messreihen ist nahezu identisch.

Die Messungen 1–8 (Tabelle 7.11) zeigen etwas größere Werte für a_N , da die gemessene THK ausgeprägter beurteilt wird: 6,5 dB (theoretisch) bzw. 6,7 dB (gemessen).

Der Mittelwert für a_N über die drei Messungen (d.h. Tabellen 7.9, 7.10, 7.11) ergibt sich anhand des Multiband-Energiemodells zu 6,1 dB, bzw. zu 6,2 dB für die durchgeführten Experimente.

Mit dieser Kenntnis kann die gemessene THK mit folgender Gleichung modelliert werden:

$$(7.9) \quad L_D := THK_{gem} = \Delta L_l + a_i + a_N \log_2(n).$$

Das Ergebnis der Modellierung ist in den Abbildungen 7.8, 7.9 und 7.10 den gemessenen Daten gegenüber gestellt. Dabei wird das Modell zum einen aus dem theoretisch ermittelten Term $a_{N_{theo}}$ (6,1 dB, „@Schwelle_Theo“) und zum anderen für den gemessenen Term $a_{N_{gem}}$ (6,2 dB, „Schwelle_gem“) visualisiert.

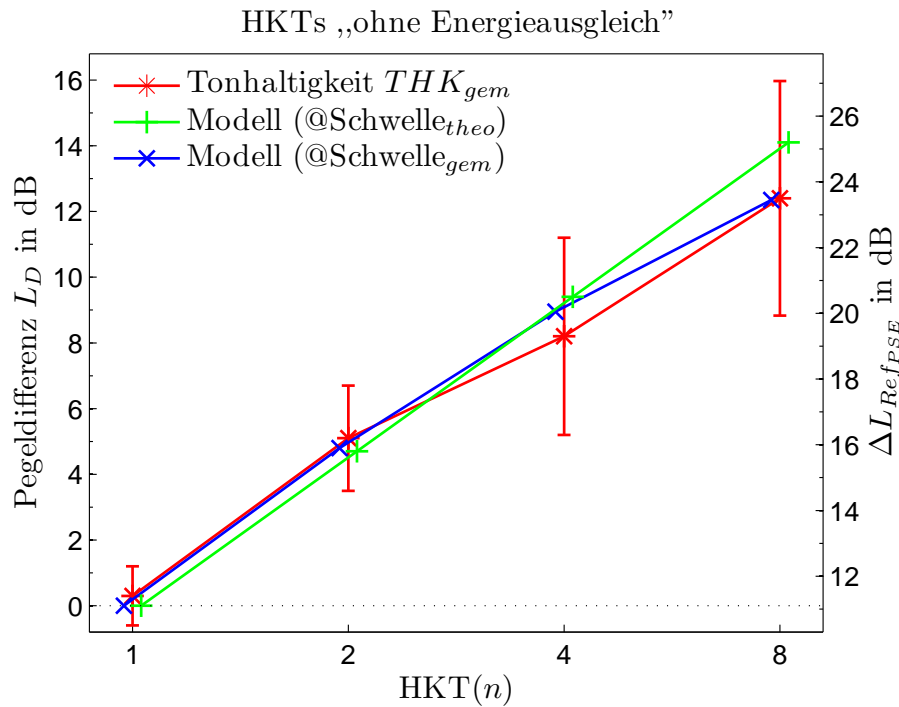


Abbildung 7.8.: Gemessene und modellierte Daten für die HKTs „ohne Energieausgleich“, vgl. Abbildung 6.3. Die drei Kurvenzüge sind zur besseren Übersicht horizontal gegeneinander verschoben.

Sowohl für die Messung der THK von HKTs „ohne Energieausgleich“ (Abbildung 7.8) als auch für die mit Blocked Tracks gemessene THK (Abbildung 7.9) wird die gemessene THK sehr gut durch die modellierten Daten beschrieben. In beiden Fällen ist die Vorhersage der THK anhand der gemessenen Mithörschwellen besser als anhand der durch das Multiband-Energiemodell vorhergesagten Mithörschwellen.

Die Modellierung der Daten für die Messungen 1–8 ist etwas schlechter als für die zuvor diskutierten Messungen, siehe Abbildung 7.10. Für die Messungen der einzelnen Tracks stimmt die Vorhersage besser als für die interleaved Tracks, grundsätzlich wird die gemessene THK durch das Modell allerdings unterschätzt. Diese Abweichung ist anhand der Tabelle 7.11 dadurch zu erklären, dass die gemessene THK in diesem Fall größer ist als in den beiden anderen Messreihen und b_N dementsprechend größer gewählt werden muss um den Fit zu verbessern.

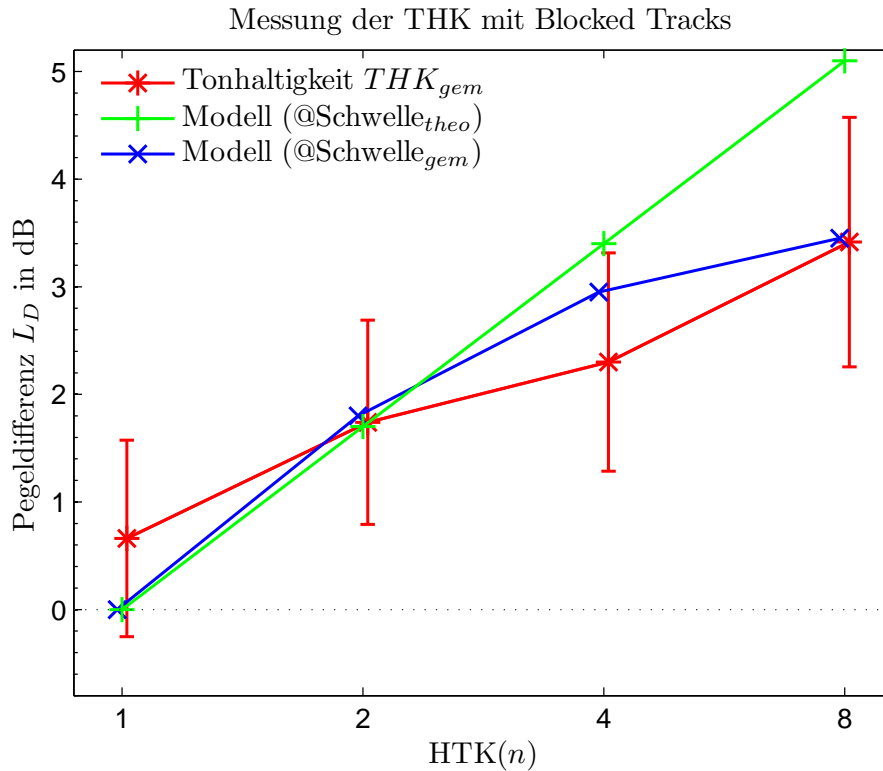


Abbildung 7.9.: Gemessene und modellierte Daten für die Blocked Tracks, vgl. Abbildung 7.7. Die drei Kurvenzüge sind zur besseren Übersicht horizontal gegeneinander verschoben.

7.8. Diskussion

Die durchgeführten Messungen der THK von HKTs zeigen konsistente Daten. Sie können gut modelliert werden, indem die THK als Einzeltonpegel über der Mithörschwelle betrachtet wird, der um die Informations-Addition a_i und einen zusätzlichen Term a_N , der die Summation mehrerer Teiltöne berücksichtigt, korrigiert wird.

$$(7.10) \quad THK = \Delta L_l + a_i + \log_2(n)a_N.$$

Die gemessene THK ist größer als es aufgrund einer Energiesummation durch hinzufügen zusätzlicher Teiltöne zu erwarten wäre (+3 dB pro Verdopplung der Teiltöne). Andererseits ist der gefundene Effekt kleiner, als es aufgrund einer Lautheitsaddition der Teiltöne vorhergesagt wird (+10 dB pro Verdopplung der Teiltöne).

Die zweite Aussage ist allerdings nur insoweit richtig, dass die erwarteten 10 dB bei Verdopplung der Lautheit für deutlich überschwellige Töne oberhalb von ≈ 40 Phon und in Ruhe gelten [FZ07]. Für Lautstärken unterhalb von

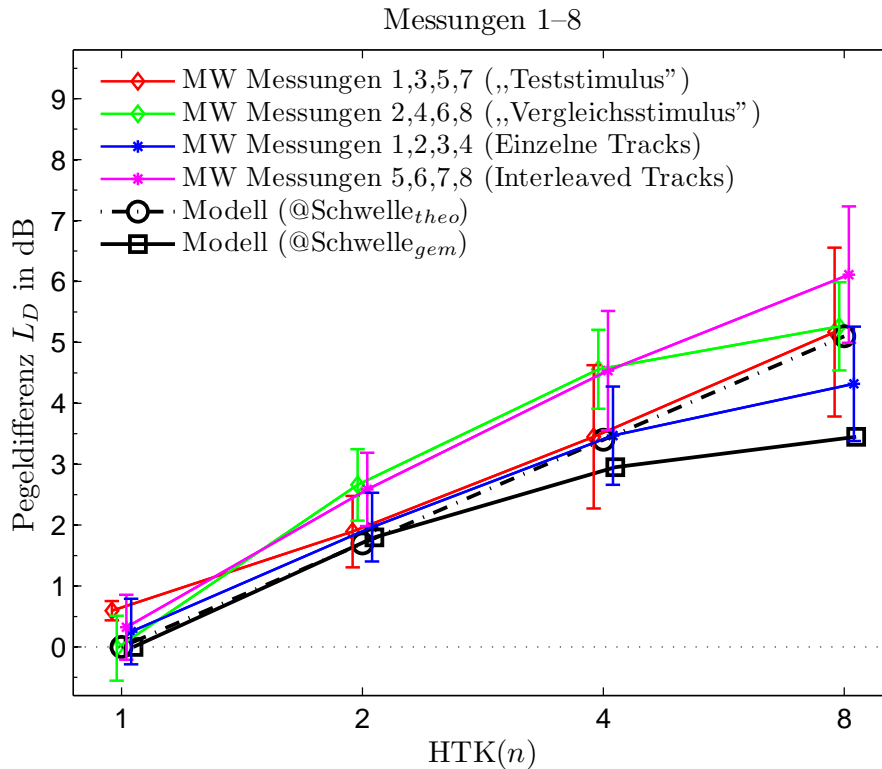


Abbildung 7.10.: Gemessene und modellierte Daten für die HKTs, Messungen 1 bis 8, vgl. Abbildung 7.6. Die drei Kurvenzüge sind zur besseren Übersicht horizontal gegeneinander verschoben.

40 Phon sind die für eine Lautheitsverdopplung erforderlichen Pegelunterschiede kleiner als 10 dB.

Um eine erste Abschätzung vorzunehmen, ob die THK sich so verhält wie es aufgrund einer Lautheitsaddition zu erwarten wäre, können die vorliegenden Ergebnisse mit denen von Fastl und Zwicker [FZ07] verglichen werden. In ihrer Abbildung 8.10 diskutieren sie eine Lautheitsfunktion für Töne in Rauschen.

Die Lautheitsfunktion eines gedrosselten Sinustones (1000 Hz) in rosa Rauschen (40 dB pro Terz) beginnt etwa bei einem Tonpegel von 37 dB. Unterhalb davon ist der Ton komplett maskiert, die Lautheit also 0 sone. Bei Erhöhung des Tonpegels steigt die Lautheit zunächst sehr steil an, ab einem Tonpegel von etwa 65 dB geht die gedrosselte in die ungedrosselte Lautheitsfunktion für den Ton über (die Steigung beträgt in beiden Fällen 10 dB pro Lautheitsverdopplung).

Für den in den vorliegenden Untersuchungen verwendeten Tonpegel von 50 dB wird aus der Abbildung in etwa eine Steigung von 5,8 dB pro Lautheitsverdopplung entnommen. Dieser Wert stimmt gut mit dem zuvor bestimmten Wert von $a_N=6,2$ dB überein.

Die Abweichungen können damit erklärt werden, dass die von Fastl und Zwicker verwendete Tonfrequenz (1000 Hz) etwas größer ist als die in dieser Studie verwendete (700 Hz). Weiterhin ist der Frequenzgruppenpegel des Maskierers hier etwas kleiner (36,2 dB) als bei Fastl und Zwicker (40 dB⁶⁹).

Scharf [Sch61] untersucht die Abhängigkeit des Abstandes der Teiltöne eines Viertonkomplexes auf die gedrosselte Lautheit des Viertonkomplexes. Er benutzt ein ähnliches Verfahren wie das hier verwendete, allerdings sind die untersuchten Frequenzen des Viertonkomplexes entweder deutlich geringer (250 Hz) oder deutlich höher (2000 Hz) als in der vorliegenden Untersuchung. Zudem sind die verwendeten Pegel der beiden Studien nur durch Interpolation ineinander überführbar.

Trotzdem kann ein grober Vergleich der beiden Studien stattfinden. Dazu wird der von Scharf gemessene Sensation Level L_{SL} seiner Abbildung 2 (Diagramm B) für den Noise-Level 60 dB und Abbildung 2 (Diagramm C) für den Noise-Level 40 dB gemittelt. Der grob abgeschätzte Sensation Level für einen Abstand der Frequenzen des Viertonkomplexes von 700 Hz liegt etwa zwischen 20 dB SL und 25 dB SL.

Der Sensation Level L_{SL} , also der Pegel oberhalb der Mithörschwelle für den „Standard-Ton“, kann mittels des o.g. Modells, Gleichung 7.10 geschätzt werden als

$$(7.11) \quad L_{SL} \approx THK = \Delta L + a_i + a_N \log_2(n).$$

Mit der maßgeblichen Differenz ΔL für den „Standard-Ton“ $\Delta L = L_T - L_G - a_v$ ergibt sich daraus

$$(7.12) \quad L_{SL} \approx L_T - L_G - a_v + a_i + a_N \log_2(n).$$

Da Tonpegel L_T und absoluter Pegel des Rauschens gleich sind, ergibt sich $L_T - L_G$ als $10 \log_{10}(24) \text{ dB} = 13,8 \text{ dB}$. Das Verdeckungsmaß für den „Standard-Ton“ beträgt $a_v(2000 \text{ Hz}) = -3,5 \text{ dB}$. Für $n=4$ und unter Verwendung des Multi-band-Energiemodells $a_i = -3 \text{ dB}$, $a_N = 6,1 \text{ dB}$ berechnet sich der Sensation-Level L_{SL} zu

$$(7.13) \quad L_{SL} \approx \Delta L = 13,8 \text{ dB} + 3,5 \text{ dB} - 3 \text{ dB} + 2 \cdot 6,1 \text{ dB} = 26,5 \text{ dB}.$$

Die Übereinstimmung der beiden Werte ist nicht perfekt, aufgrund der nur indirekten Vergleichbarkeit der beiden Studien aber ausreichend.

⁶⁹ Der angegebene Pegel gilt für eine Bandbreite von einer Terz, d.h. $\approx 260 \text{ Hz}$ bei der Tonfrequenz von 1000 Hz. Die Bandbreite der Frequenzgruppe um 1000 Hz beträgt nach [FZ07] hingegen 160 Hz. Dadurch ist der Pegel pro Frequenzgruppe tatsächlich etwas kleiner als die angegebenen 40 dB.

Die diskutierten Vergleiche weisen darauf hin, dass die vorliegenden Daten zur THK im Wesentlichen durch die Lautheit bestimmt werden. Diese Feststellung deckt sich mit dem Befund von Hansen [Han10], der die Effekte der THK ebenfalls auf die Lautheit zurückführt.

Die hier gefundenen Beiträge zur THK aufgrund der (gedrosselten) Lautheitsaddition a_N gelten nur für maßgebliche Differenzen, die auch während der Messungen verwendet wurden, also etwa zwischen 14 dB und 20 dB liegen. Es bleibt zu prüfen, ob das Modell erweitert werden kann, indem für andere maßgebliche Differenzen der Beitrag zur Lautheitsaddition a_N durch eine Lautheitsfunktion für (gedrosselte) Töne in Rauschen approximiert werden kann (vgl. hierzu z.B. Abbildung 8.10 aus Fastl und Zwicker [FZ07]).

Bezüglich der Evaluation der Verfahren bleibt festzuhalten, dass das adaptive THK-Messverfahren einfach durchzuführen ist und reproduzierbare und konsistente Messwerte liefert.

Die Messungen können zeitlich optimiert werden, indem ein zufälliger Startwert bei der Messung gewählt wird. Ein Unterschied zu Messungen, bei denen der Startpegel systematisch variiert und über mehrere einzelne Messungen (Tracks) gemittelt wird, kann nicht nachgewiesen werden.

Das Messverfahren mit Interleaved Tracks ist als konservativer einzuschätzen, da die THK mit diesem Messdesign eher größer eingeschätzt wird als mit den anderen getesteten Messdesigns. Im Sinne der Norm wird somit die im Zweifel zu große THK zugunsten von geringeren Geräuschemissionen gemessen.

Einzelne Tracks sind zu empfehlen, wenn die THK nur für einen einzigen Parameter gemessen oder eine schnelle Messung durchgeführt werden soll, für die eine etwas gröbere Auflösung ausreichend sind.

Die Messungen mit dem Messdesign Blocked Tracks sind für die Anwendung in der gutachterlichen Praxis nicht geeignet. Für die Grundlagenforschung zur THK ist das Verfahren hingegen brauchbar und wertvoll.

7. *Verfahrensvergleich zur Bestimmung der Tonhaltigkeit*

8. Zusammenfassung

Im Rahmen dieser Arbeit wurde die Wahrnehmung der Tonhaltigkeit untersucht. Dabei zeigte sich, dass der Tonhaltigkeit eine interne Repräsentation (Wahrnehmungsraum) des Wahrnehmenden zugrunde liegt, die für alle Individuen nahezu gleiche Muster und Strukturen aufweist.

Gleichwohl divergieren die individuellen verbalen Beschreibungen der Tonhaltigkeit und können aufgrund ihrer Komplexität nur nach aufwendiger Kodierung ineinander überführt und verglichen werden. Da kein einheitlicher Wortschatz bzgl. der Tonhaltigkeit existiert, sind Unklarheiten und Mehrdeutigkeiten die Konsequenz, die eine subjektive Skalierung der Tonhaltigkeit erschweren.

Um diese Schwierigkeit zu überwinden, kann eine akustische Demonstration der Tonhaltigkeit (Orientierung) vorgenommen werden. Auf diese Weise wird der Wahrnehmung „Tonhaltigkeit“ ein geeignetes Vokabular zugewiesen. Ein Vorschlag zur Orientierung wurde in der vorliegenden Arbeit entwickelt, mit dem diese Aufgabe angemessen durchgeführt werden kann.

Ein universell anwendbares, adaptives Verfahren zur Bewertung (Messung) der wahrgenommenen Tonhaltigkeit wurde erarbeitet, eingeführt und evaluiert. Das Ergebnis dieser Messungen sind Pegelwerte, die leicht zu interpretieren sind und daher unmittelbar quantitative Aussagen ermöglichen.

Die Effizienz des Verfahrens konnte anhand der durchgeführten Messungen validiert werden. Schon mit einer geringen Teilnehmerzahl während der Messungen konnten statistisch belastbare und valide Daten ermittelt werden.

Explorative Messungen verschiedener Parameter der Tonhaltigkeit wurden durchgeführt. Die Ergebnisse bestätigen zum Teil die Kriterien und Voraussetzungen der aktuellen Norm zur Tonhaltigkeit DIN 45681:2005-03. Einige in der Norm verwendete Annahmen konnten jedoch nicht vollständig verifiziert werden.

Dazu gehört die Frequenzabhängigkeit der Tonhaltigkeit. Sie wird, wie in der DIN 45681 vorgesehen, für Frequenzen oberhalb von 500 Hz angemessen durch die Mithörschwelle berücksichtigt. Unterhalb dieser Frequenz findet hingegen eine Überschätzung der Tonhaltigkeit statt.

Dieser Befund wird durch zusätzliche Messungen mit Paarvergleichen verifiziert und kann durch ein Modell der Ausgeprägtheit von Tönen („Pitch Strength“) erklärt werden.

Die Tonhaltigkeit von Schmalbandrauschen in einem breitbandigen Hintergrund wird durch die DIN 45681 für kleine Bandbreiten angemessen berücksichtigt, für größere Bandbreiten jedoch unterschätzt.

Die Integration mehrerer tonaler Komponenten erfolgt in der DIN 45681 durch Verwendung der am stärksten ausgeprägten Komponente unter Nichtbeachtung der anderen Komponenten. Diese Verfahrensweise unterschätzt die Tonhaltigkeit bei gleichzeitiger Anwesenheit von mehreren, vergleichbar ausgeprägten Komponenten erheblich.

Die Tonhaltigkeit kann für diesen Fall adäquat modelliert werden, indem die gedrosselten Lautheiten der einzelnen tonalen Komponenten addiert werden. Die Addition hängt bezüglich der zugrunde liegenden Pegelwerte der tonalen Komponenten von der Steigung der nichtlinearen (gedrosselten) Lautheitsfunktion ab. In der vorliegenden Arbeit wurde dieses Modell für den in der Arbeit betrachteten mittleren Bereich der Tonhaltigkeit entwickelt.

Alle hier diskutierten Parameter der Tonhaltigkeit weisen darauf hin, dass die Tonhaltigkeit im Wesentlichen durch die (gedrosselte) Lautheit bestimmt wird.

A. Anhang: breitbandiges Rauschen, Verdeckung und Verdeckungsmaß

A.1. Breitbandiges Rauschen

In dieser Arbeit werden die nachfolgend aufgelisteten breitbandigen Rauschtypen behandelt bzw. verwendet. Ihre wichtigsten Eigenschaften werden im Weiteren diskutiert.

- weißes Rauschen
- rosa Rauschen ($1/f$ -Rauschen)
- braunes Rauschen ($1/f^2$ -Rauschen)
- mittlerer (Rausch-)Hintergrund von Umweltschallen
- gleichmäßig anregendes Rauschen (GAR)
- gleichmäßig verdeckendes Rauschen (GVR)

Weißes Rauschen

Weißes Rauschen wird im Wesentlichen durch eine konstante Rauschleistungsdichte im Leistungsdichtespektrum charakterisiert, vgl. Abbildung [A.1](#).

Rosa Rauschen

Die Rauschleistungsdichte für rosa Rauschen halbiert sich bei Verdopplung der Frequenz (der Abfall beträgt 3 dB/Oktave). Daher wird rosa Rauschen auch als $1/f$ -Rauschen bezeichnet.

Als Faustregel gilt, dass rosa Rauschen alle Frequenzbereiche des hörbaren Schallspektrums oberhalb von ca. 500 Hz in etwa gleich anregt.

Braunes Rauschen

Für braunes Rauschen fällt die Rauschleistungsdichte um 6 dB/Oktave, d.h. die Rauschleistungsdichte verringert sich bei Verdopplung der Frequenz auf ein Viertel. Braunes Rauschen wird auch als $1/f^2$ -Rauschen bezeichnet. Im englischsprachigen Raum ist die Notation „rotes Rauschen“ gebräuchlich.

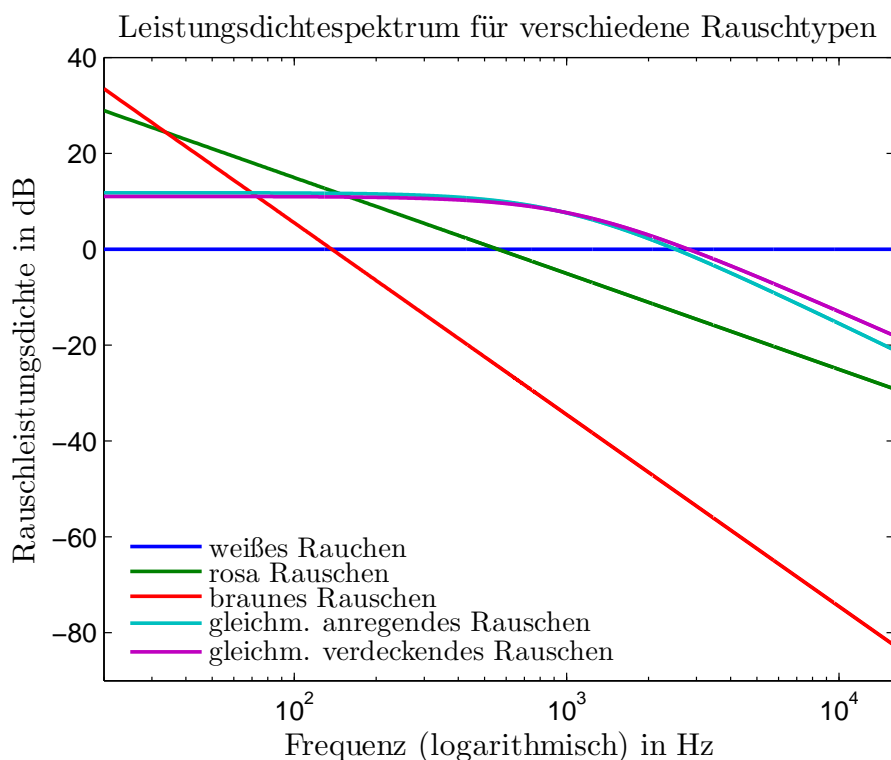


Abbildung A.1.: Leistungsdichtespektrum für verschiedene Rauschtypen über der Frequenz. Aufgetragen ist der relevante Hörbereich des Menschen. Alle Rauschen weisen die gleiche Leistung (RMS-Wert) auf, die Normierung erfolgt auf das weiße Rauschen.

Braunes Rauschen wirkt gegenüber den beiden erst genannten Rauschen weniger scharf und wird daher als angenehmer empfunden.

Mittlerer (Rausch-)Hintergrund von Umweltschallen

Zur Untersuchung der THK von realen Umweltschallen kann ein Rauschen erzeugt werden, welches dem mittleren Spektrum von typischen Umweltschallen entspricht. Als repräsentative Umweltschalle werden hier die Umweltgeräusche von Pompetzki [Pom98], siehe Anhang B.1 verwendet. Das mittlere Spektrum dieser Geräusche hat für Frequenzen unterhalb von ca. 2000 Hz in etwa den Verlauf des braunen Rauschens, d.h. der Abfall der Rauschleistungsdichte beträgt etwa 6 dB/Oktave. Für Frequenzen oberhalb von ca. 2000 Hz wird der Abfall steiler und beträgt dort um 12 dB/Oktave.

Soll allgemein ein prototypisches Rauschen für Umweltschalle verwendet werden, so kann am ehesten das braune Rauschen als Näherung dienen.

Gleichmäßig anregendes Rauschen (GAR)

Die drei erstgenannten Rauscharten sind eher technisch bedingt, während die beiden folgenden aus psychoakustischen Überlegungen motiviert sind.

Gleichmäßig anregendes Rauschen (engl. „uniformexciting noise“) ist so definiert, dass es eine gleichmäßige, d.h. konstante Erregung auf der Basilarmembran hervorruft und daher in jeden auditorischen Filter die gleiche Energie fällt. Die Schallintensität für GAR ist somit in jeder Frequenzgruppe identisch und der Frequenzgruppenpegel für alle Mittenfrequenzen konstant.

Gleichmäßig verdeckendes Rauschen

Gleichmäßig verdeckendes Rauschen (engl. „uniformmasking noise“) verdeckt konstant präsentierte Sinustöne (Dauertöne) für alle Frequenzen im Hörbereich gleichmäßig, d.h. die Mithörschwelle für diese Töne ist in gleichmäßig verdeckendem Rauschen über den gesamten Frequenzbereich konstant.

Die frequenzabhängige Schallintensitätsdichte für gleichmäßig verdeckendes Rauschen ist unterhalb von 500 Hz konstant und fällt darüber näherungsweise mit etwa 10 dB/Dekade ab [Zwi82].

A.2. Verdeckungsmaß

Das *Verdeckungsmaß* beschreibt den Unterschied zwischen GAR und GVR. Der Unterschied wird hervorgerufen durch die Frequenzabhängigkeit der Frequenzgruppenbreite: Für Mittenfrequenzen f_c der Frequenzgruppe unterhalb 500 Hz beträgt die Frequenzgruppenbreite Δf_c konstant etwa 100 Hz, darüber hinaus steigt Δf_c etwa proportional einer kleinen Terz (entspricht dem Faktor 1,19) an.

Bei hohen Frequenzen kann das Rauschen in einer Frequenzgruppe aufgrund der deutlich größeren Frequenzgruppenbreite als nahezu stationär angesehen werden, während das Rauschen bei tiefen Frequenzen stärker zeitlich fluktuiert. Dieser Effekt wird durch die Eigenmodulation (engl. „self-modulation“) des Rauschens hervorgerufen, vgl. Fastl und Zwicker [FZ07], Seite 199.

Die Eigenmodulation führt dazu, dass tieffrequente Töne bei Maskierung durch gleichmäßig anregendes Rauschen weniger gut wahrnehmbar sind als Töne mit höheren Frequenzen.

Das Verdeckungsmaß a_v ist psychoakustisch definiert als $a_v = L_T - L_G$. Dabei beschreibt L_G den Pegel in der Frequenzgruppe mit der Mittenfrequenz f_c um den maskierten Ton mit der Frequenz f_T und dem Pegel L_T .

Quantitativ wird das Verdeckungsmaß durch die folgende Formel [THK06]

bestimmt:

$$(A.1) \quad a_v = \left[-2 - \log \left(1 + \left(\frac{f}{502 \text{ [Hz]}} \right)^{2.5} \right) \right] \text{ dB.}$$

Der frequenzabhängige Verlauf über die Mittenfrequenz f_c der verdeckenden Frequenzgruppe $a_v(f_c)$ ist in Abbildung A.2 aufgetragen. Für hohe Frequenzen ist a_v etwa -6 dB, für tiefe Frequenzen -2 dB. Fastl und Zwicker [FZ07] diskutieren diesen Sachverhalt eingehend.

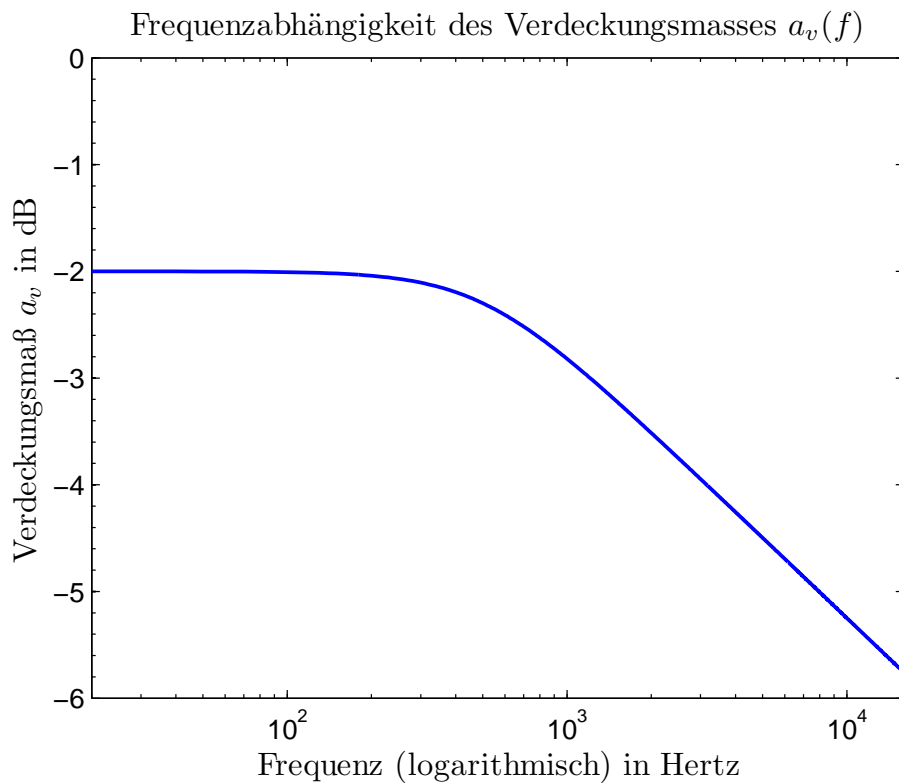


Abbildung A.2.: Verdeckungsmaß $a_v(f)$ nach Gleichung A.1 über der Frequenz. Aufgetragen ist der relevante menschliche Hörbereich.

B. Anhang: Geräusch-Sets und Signale

B.1. Pompetzki

Pompetzki vom Landesumweltamt Essen untersucht 1998 vierzehn Geräusche hinsichtlich der Bewertung der THK [Pom98]. Es wird die subjektive Beurteilung durch Gutachter⁷⁰ mit den algorithmisch berechneten Ergebnissen der DIN 45681:1992-01e [THK92] verglichen.

Tabelle B.1 enthält eine Übersicht der verwendeten Geräusche, die die gängige Praxis im Immissionsschutz repräsentieren. Die Geräusche enthalten (typische) tonale Komponenten zwischen 109 Hz und 1008 Hz.

Tabelle B.1.: Umweltgeräusche von Pompetzki mit den von ihm anhand der DIN 45681:1992-01e [THK92] berechneten Tonfrequenzen (d.h. die Frequenz mit dem höchsten Tonzuschlag K_T), den maßgeblichen Differenzen ΔL und den Tonzuschlägen K_T . Die angegebenen Werte wurden mir freundlicherweise von Herrn Pompetzki zur Verfügung gestellt.

Nr.	Signal-Art	Tonfrequenz	ΔL	K_T
LUA 1	Glashütte	299 Hz	12,2 dB	6 dB
LUA 2	Pumpenanlage	629 Hz	10,3 dB	6 dB
LUA 3	Hydraulikpumpe	400 Hz	17,1 dB	6 dB
LUA 4	WKA normal	190 Hz	4,7 dB	3 dB
LUA 5	Schreinerei	578 Hz	6,1 dB	4 dB
LUA 6	Papierpresse	223 Hz	11,5 dB	6 dB
LUA 8	Kühlaggregat	109 Hz	13,3 dB	6 dB
LUA 9	WKA defekt	1008 Hz	10,2 dB	6 dB
LUA 10	Industrietrafo	109 Hz	19,5 dB	6 dB
LUA 11	Großchemie	494 Hz	10,4 dB	6 dB
LUA 12	Lüftungsanlage	129 Hz	4,8 dB	3 dB
LUA 14	Wandlüfter	114 Hz	4,3 dB	3 dB

⁷⁰ so wie in TA-Lärm von 1998 vorgeschrieben

B. Geräusch-Sets und Signale

Das Signal des Autoscooters (LUA 7) und des Glockenspiels (LUA 13) sind hier aus den von Pompetzki diskutierten Gründen [Pom98] weggelassen.

C. Anhang: Perzeptuelle Aspekte der Tonhaltigkeit

Im Folgenden werden die in Matlab implementierten Benutzer-Oberflächen (graphical user interface, GUI) zur computergesteuerten Durchführung des RepGrid-Verfahrens vorgestellt. Für Einzelheiten zum Verfahren vergleiche Kapitel 3.

C.1. Evokationsphase

Die Oberfläche für die Evokationsphase ist in Abbildung C.1 dargestellt. Die drei grünen Buttons 1, 2, 3 im oberen linken Bereich sind mit den Stimuli einer Triade in zufälliger Reihenfolge belegt.

Zu Beginn der Untersuchung werden dem Probanden die drei Stimuli nacheinander präsentiert. Im Folgenden kann er die Stimuli einzeln durch Anklicken der Buttons 1, 2, 3 wiederholen, oder bei Benutzung der Schaltfläche **Darbietung wiederholen** die komplette Präsentation der Triade neu starten.

Die Urteilsabgabe (welche zwei Stimuli sich ähnlich sind) erfolgt im mittleren Bereich durch Betätigen einer der drei blauen Schaltflächen **1 u.2**, **1 u.3**, **2 u.3**. Die Antwort wird dem Probanden durch Anzeige der entsprechenden Beurteilung im hellroten Feld im Bereich **Antwort** bestätigt.

Anschließend muss der Proband über die Tastatur mindestens jeweils eine Eigenschaft (=Konstrukt) für Gemeinsamkeiten bzw. Unterschiede der Stimuli in die entsprechenden Felder eingeben. Erst danach kann er durch Betätigen der blauen Schaltfläche **Fertig** oben rechts die Bewertung dieser Triade abschließen und gelangt automatisch zur Bewertung der nächsten Triade.

Die Eintragungen in Abbildung C.1 sind willkürlich und haben keinen Bezug zu den durchgeführten Experimenten.

C.2. Beurteilungsphase

Es finden zwei verschiedene Methoden Anwendung, die Elemente (akustische Stimuli) bzgl. der Konstrukte zu bewerten. Zum einen wird mittels der Benutzeroberfläche aus Abbildung C.2 eine Rangordnung der Elemente hergestellt. Zum anderen werden die Elemente auf einer quasikontinuierlichen Skala (Abbildung C.3) bewertet.

Abbildung C.1.: Repertory-Grid: Benutzeroberfläche zur Durchführung der Evokationsphase, in der die Konstrukte für die Elemente (akustische Stimuli) erhoben werden

Bei der Herstellung der Rangreihe werden dem Probanden zunächst alle sechs Stimuli A – F einmal präsentiert. Durch Betätigen der entsprechenden Schaltflächen A – F können die Schallsignale einzeln jederzeit und beliebig oft wiederholt werden (**Darbietung wiederholen** spielt alle Schallsignale erneut nacheinander ab).

Die gelben Schaltflächen sind zu Anfang nicht belegt und werden durch Verwendung der sechs vertikal angeordneten **Drop-Down Comboboxen** durch den Probanden nacheinander entsprechend der in den grauen Textfeldern angegebenen Kontrastpole (hier tonhaltig – nicht tonhaltig) zugewiesen.

Nach der Zuweisung können die Schallsignale auch dadurch abgespielt werden, dass die gelben Schaltflächen A – F benutzt werden. Das ist nützlich, um die Reihenfolge, in der die Stimuli angeordnet werden, abschließend zu überprüfen.

Der Proband schließt die Bewertung durch die Schaltfläche **Fertig** ab. In diesem Moment überprüft die Benutzeroberfläche, ob jeder Stimulus genau einmal zugewiesen worden ist. Dieser Prozess wird sukzessive für alle Konstrukte bzw.

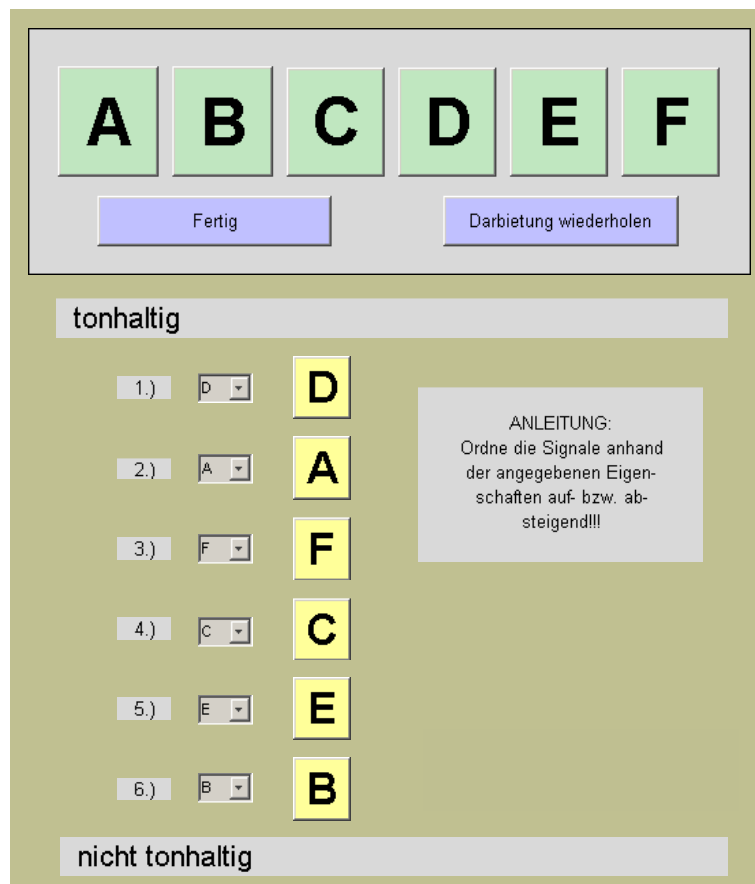


Abbildung C.2.: Repertory-Grid: Benutzeroberfläche zur Bewertung der akustischen Stimuli (Elemente) bzgl. der Konstrukte durch Herstellung einer Rangreihe

Kontrastpole nacheinander durchgeführt.

Bei der Beurteilung auf der quasikontinuierlichen Skala ist eine verkürzte Version der Anleitung für den Probanden im oberen, linken Bereich **Anleitung** jederzeit sichtbar.

Zu Beginn der Skalierung wird der zu bewertende Stimulus einmal dem Probanden präsentiert. Durch Betätigen der Schaltfläche **Darbietung wiederholen** kann das Schallsignal jederzeit und beliebig oft erneut abgespielt werden.

Im unteren Bereich sind zwischen einer und fünf Skalen angeordnet. Ihre Pole sind durch die jeweiligen Kontrastpaare (Konstrukte) aus der Evokationsphase gegeben.

Zur Bewertung wird ein **Schieber** verwendet, den die Probanden quasikontinuierlich einstellen können⁷¹.

⁷¹ Effektiv sind 101 Stufen zwischen 0 und 100 mit dem Schieber realisierbar.

C. Perzeptuelle Aspekte der Tonhaltigkeit

The image shows a software interface for a Repertory-Grid experiment. It is contained within a light green border. At the top left, a grey box contains the following instructions in German: "ANLEITUNG: 1.) Höre die Darbietung an, 2.) Stelle die Schieber ein!, 3.) Bestätige die Einstellungen!, 4.) --> 'Nächste Darbietung' am Ende: --> 'Schliessen'". To the right of this box are two blue buttons: "Fertig" (Finished) and "Darbietung wiederholen" (Repeat presentation). Below the instructions are three horizontal sliders. The first slider is labeled "tonhaltig" (tonal) on the left and "nicht tonhaltig" (not tonal) on the right, with a slider knob positioned approximately one-third of the way from the left. The second slider is labeled "Rauschen" (noise) on the left and "Ton" (tone) on the right, with a slider knob positioned approximately two-thirds of the way from the left. The third slider is labeled "Ton leise" (soft tone) on the left and "Ton laut" (loud tone) on the right, with a slider knob positioned approximately two-thirds of the way from the left. Below these sliders are two empty grey rectangular boxes for notes. At the bottom of the interface is a checkbox labeled "Bestätigung: Alle beschrifteten Schieber eingestellt!" (Confirmation: All labeled sliders set!).

Abbildung C.3.: Repertory-Grid: Oberfläche zur Bewertung der Schallsignale (Elemente) bzgl. der Konstrukte mittels einer quasikontinuierlichen Skala (=„Schieber“)

Durch Anhaken des Feldes **Alle beschrifteten Schieber eingestellt** bestätigt der Proband, dass er alle Bewertungen abgeschlossen hat. Erst danach kann er die Schaltfläche **Fertig** zum Beenden der Skalierung dieses Stimulus benutzen.

Die Bewertung wird auf diese Weise sukzessive für alle Stimuli durchgeführt.

D. Anhang: Wahl eines Orientierungsschalls für die THK

D.1. Orientierungsschalle

Die potentiellen Orientierungsschalle werden mit einer Lautheit von $10 \text{ sone}_{\text{GF}}$ dargeboten. Zur Kalibrierung der Lautstärke während der Darbietung eignen sich Angaben bezüglich des Lautstärkepegels jedoch besser. Aus diesem Grund werden die Lautheiten nach DIN 45631 / ISO 532B für das Freifeld bei frontalem Einfall für diotische Darbietung mittels der SIP-Toolbox [Fra10] in äquivalente Dauerschallpegel L_{EQ} umgerechnet, siehe Tabelle D.1.

Tabelle D.1.: Äquivalenter Dauerschallpegel L_{EQ} der Orientierungsschalle für die Signaldauer von 30 s.

Potentieller Orientierungsschall	Lautheit in sone_{GF}	Äquivalenter Dauerschallpegel in dB SPL	Äquivalenter Dauerschallpegel in dB(A) SPL
Schall 1	10,0	59,7	57,7
Schall 2	10,0	61,1	59,3
Schall 3	10,0	63,6	60,2
Schall 4	10,0	59,8	57,8

Durch Anklicken der Schalle in Tabelle D.1 wird bei geeigneter Konfiguration und verfügbarem Zugriff auf das Internet der jeweilige Schall abgespielt. Alternativ können die Schalle auch direkt von der Webseite www.tonhaltigkeit.de heruntergeladen oder angehört werden.

D.2. Tonhaltigkeitsverlauf der Orientierungsschalle

Der Verlauf der Tonhaltigkeit für die vier potentiellen Orientierungsschalle kann mit der Norm DIN 45681:2006-08 [THK06] nicht direkt berechnet werden, da sich der Pegel der tonalen Komponente zu schnell ändert, vgl. Kapitel 2.10.

D. Wahl eines Orientierungsschalls für die THK

Aus diesem Grund werden die Parameter, mit der die THK berechnet wird, gegenüber der Norm angepasst.

Die Orientierungsschalle liegen mit einer Samplingfrequenz von 48 kHz vor. Bei einer Blockgröße von 16384 Samples beträgt die Frequenzauflösung Δf hier 2,930 Hz und die Zeitauflösung Δt 0,341 s. Um die Zeitverläufe der Orientierungsschalle abbilden zu können, wird hier nur über drei Spektren gemittelt. Die Analyse ist somit streng genommen nicht normgerecht, für diesen Zweck aber zielführend und ausreichend.

Die auf diese Weise über etwa 1 s gemittelten maßgeblichen Differenzen ΔL sind in Abbildung D.1 dargestellt.

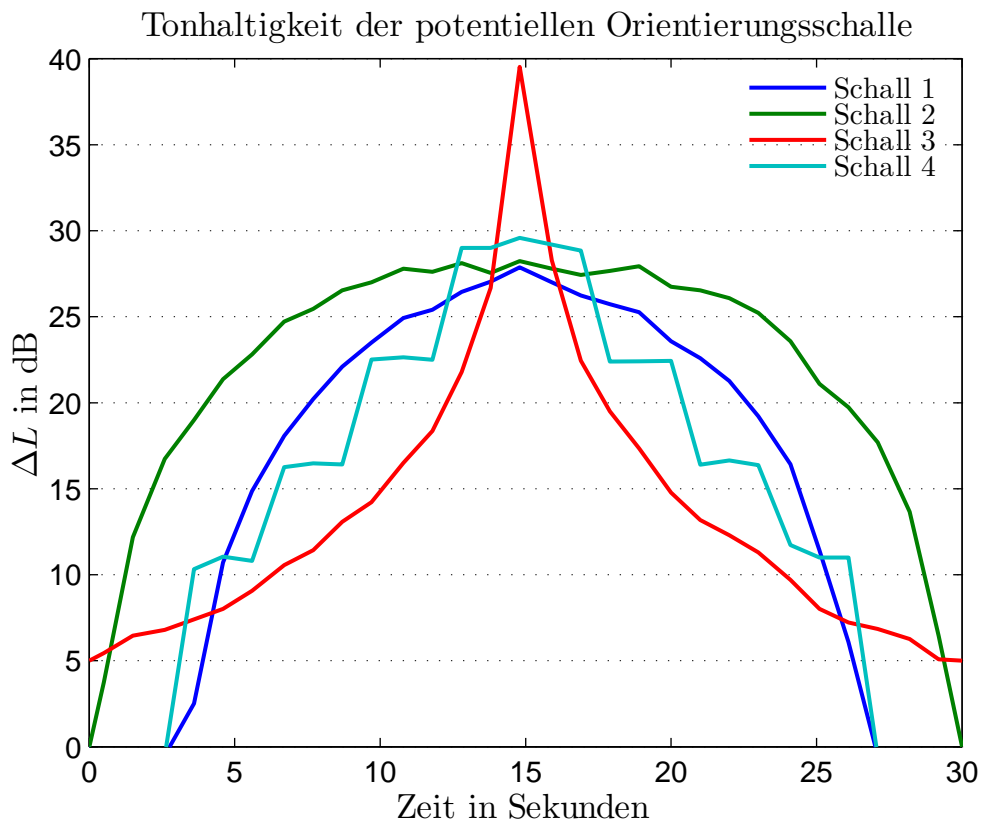


Abbildung D.1.: Zeitlicher Verlauf der THK (maßgebliche Differenz, ΔL) für die potentiellen Orientierungsschalle.

Im Gegensatz zu den schematisch dargestellten Zeitverläufen aus Abbildung 4.1 handelt es sich in Abbildung D.1 um die tatsächlich vorliegenden Werte der THK für die vier potentiellen Orientierungsschalle.

D.3. Individuelle Probandenurteile

Die von den Probanden abgegebenen Urteile sind in Tabelle D.2 zur Übersicht aufgelistet und zusammengefasst.

Tabelle D.2.: Kategoriale Urteile für die drei Probandengruppen. Die Reihenfolge der Probanden ist zufällig: Für die beiden Experten-Gruppen sind diejenigen Probanden, die an beiden Untersuchungen teilgenommen haben, nicht miteinander verbunden, d.h. in der gleichen Zeile gelistet. Probanden ohne eindeutige Präferenz (siehe Abschnitt 4.6) sind hier grau hinterlegt.

Proband	Kategoriale Urteile für											
	Experten 1: Schall-Nr.				Experten 2: Schall-Nr.				Laien: Schall-Nr.			
	1	2	3	4	1	2	3	4	1	2	3	4
1	3	3	1	2	4	3	1	1	3	3	3	3
2	3	3	3	3	1	1	1	4	2	4	1	2
3	2	2	2	3	4	4	1	2	1	1	2	2
4	4	2	3	4	3	3	4	1	3	2	3	4
5	3	3	2	3	4	3	3	4	2	3	2	1
6	2	3	3	2	4	3	2	2	1	3	2	1
7	3	3	2	2	3	4	1	3	2	3	1	2
8	4	2	2	3	3	2	1	4	3	3	1	1
9	1	1	2	2	3	3	1	2	3	3	4	4
10	3	3	1	2	3	3	2	3	2	3	2	3
11	4	3	4	4	3	3	2	3	2	4	1	3
12	3	2	1	2	3	4	2	2	2	4	1	3
13	3	3	1	1	2	3	2	4	3	3	2	2
14	3	2	1	2	3	3	1	2	1	3	2	3
15	3	1	4	2	3	3	1	2	4	4	3	4
16	2	2	1	3	3	3	3	2	2	2	3	3
17	2	3	1	3								
18	3	3	2	2								
19	3	3	3	3								
20	4	4	3	4								

D.4. Law of Categorical Judgement

Das *Law of Categorical Judgement* geht auf Thurstone (1927) zurück. Ziel dabei ist „Daten von einem niedrigeren (hier ordinalen) Skalenniveau auf ein höheres

Skalenniveau (hier Intervallskala) zu transformieren” [BD06].

Um diese Transformation durchführen zu können, müssen einige Voraussetzungen erfüllt sein, darunter vor allem die Normalverteilung der bewerteten Merkmalsausprägungen.

Andererseits werden die mit dem Law of Categorical Judgement berechneten Ergebnisse wenig durch nicht normalverteilte Bewertungen der Merkmalsausprägung beeinflusst. Darüber hinaus sind die Ergebnisse invariant gegenüber unterschiedlichen Urteilerstichproben, Kategorienbezeichnungen und selbst der Anzahl der Kategorien, siehe hierzu Bortz und Döring [BD06].

Aufgrund der Robustheit des Verfahrens wird im Folgenden die Anwendbarkeit für den vorliegenden Datensatz angenommen, ohne alle Voraussetzungen exakt zu überprüfen.

Im weiteren Verlauf wird der Rechengang nach Bortz und Döring ([BD06], Seite 158) kurz beschrieben. Um die vorliegenden Daten anhand des Law of Categorical Judgement zu transformieren, werden zunächst die Häufigkeiten der kategorialen Urteile für die Orientierungsschalle (Tabelle D.2) gruppenweise ermittelt. Die derart gewonnenen Häufigkeiten zeigt Tabelle D.3.

Tabelle D.3.: Häufigkeit der kategorialen Urteile für die drei Probandengruppen.

Schall-Nr.	Experten 1: Kategorie				Experten 2: Kategorie				Laien: Kategorie			
	1	2	3	4	1	2	3	4	1	2	3	4
1	1	4	11	4	1	1	10	4	3	7	5	1
2	2	6	11	1	1	1	11	3	1	2	9	4
3	7	6	5	2	8	5	2	1	5	6	4	1
4	1	9	7	3	2	7	3	4	3	4	6	3

Die absoluten Häufigkeiten werden nun in relative Häufigkeiten umgerechnet, indem durch die Anzahl der Urteile geteilt wird. Anschließend werden die relativen Häufigkeiten zeilenweise kumuliert, siehe Tabelle D.4.

Die kumulierten Häufigkeiten werden als Flächenanteile der Standardnormalverteilung angesehen und die korrespondierenden z -Werte (Abszissenwerte) der Standardnormalverteilung ermittelt. Diese z -Werte sind exemplarisch für die ExpertenGruppe 2 in Tabelle D.5 aufgelistet. Die Kategorie 4 ist hier nicht mehr verzeichnet, da die kumulierte relative Häufigkeit stets 1 ist und der z -Wert somit gegen unendlich konvergiert.

Die drei Kategoriengrenzen zwischen den vier verwendeten Kategorien werden durch die Spaltenmittelwerte der Tabelle D.5 repräsentiert. Die Merkmalsausprägungen, d.h. die intervallskalierten Skalenwerte für die beurteilten Orientierungsschalle, resultieren aus der Differenz der mittleren Kategoriengrenze

Tabelle D.4.: Kumulierte relative Häufigkeiten der kategorialen Urteile für die drei Probandengruppen.

Schall-Nr.	Experten 1: Kategorie				Experten 2: Kategorie				Laien: Kategorie			
	1	2	3	4	1	2	3	4	1	2	3	4
1	.05	.25	.80	1.0	.06	.13	.75	1.0	.19	.63	.94	1.0
2	.10	.40	.95	1.0	.06	.13	.81	1.0	.06	.18	.75	1.0
3	.35	.65	.90	1.0	.50	.81	.94	1.0	.31	.69	.94	1.0
4	.05	.50	.85	1.0	.13	.56	.75	1.0	.19	.44	.81	1.0

Tabelle D.5.: Abszissenwerte der Standardnormalverteilung (z -Werte), exemplarisch für die Gruppe „Experten 2“. Im rechten Teil der Tabelle sind die Zeilensummen (ZS), Zeilenmittelwerte (ZM) und die Merkmalsausprägungen (MA und MA_0) eingetragen. Im unteren Teil sind die Spaltensummen (SS) und die Spaltenmittelwerte (SM) erfasst. Die Spaltenmittelwerte repräsentieren zugleich die Kategoriengrenzen (KG).

Schall-Nr.	z -Werte für Kategorie			ZS	ZM	MA	MA_0
	1	2	3				
1	-1.53	-1.15	0.67	-2.01	-0.67	0.53	1.48
2	-1.53	-1.15	0.89	-1.80	-0.60	0.46	1.41
3	0.00	0.89	1.53	2.42	0.81	-0.95	0.00
4	-1.15	0.16	0.67	-0.32	-0.11	-0.04	0.91
SS	-4.22	-1.26	3.77				
KG($\hat{=}$ SM)	-1.05	-0.31	0.94		-0.14		

(-0.14) zu den Zeilenmittelwerten. Diese Werte stellen die gesuchte, intervallskalierte Repräsentation der Daten dar. Die entsprechenden Zahlenwerte sind in der zweiten Spalte von rechts in Tabelle D.5 eingetragen.

Es ist an dieser Stelle üblich, durch Addition des kleinsten Wertes der intervallskalierten Daten, den Nullpunkt der Skala zu verschieben (rechte Spalte (MA_0) in Tabelle D.5). Auf diese Weise können die Daten für die drei Gruppen leichter in ein gemeinsames Diagramm (Abbildung 4.6) eingetragen und interpretiert werden.

D.5. Erweiterte Analyse der Daten (ANOVA)

Unter der Annahme, dass die Skala aus Abbildung 4.2 gleiche Abstände zwischen den Kategorien aufweist, kann für die Urteile der Probanden aus Tabelle D.2 Intervallskalenniveau vorausgesetzt werden.

Unter diesen Bedingungen wird eine zweifaktorielle, univariate ANOVA mit der Statistik-Software SPSS⁷² durchgeführt. Ein Faktor kodiert die Gruppen (dreistufig), der andere die Orientierungsschalle (vierstufig).

Die ANOVA deckt keinen Effekt ($F(2,196) < 1$) der Gruppe auf die Bewertungen der potentiellen Orientierungsschalle auf. Die Schalle selber haben hingegen einen signifikanten Effekt ($F(3,196) = 10.23$, $p < .001$). Die Interaktion zwischen diesen beiden Faktoren (Gruppe*Orientierungsschall) ist lediglich tendenziell signifikant ($F(6,196) = 2.00$, $p = .067$).

Die Effektgröße, d.h. die Varianzaufklärung für den Faktor „Orientierungsschalle“, liegt nach Bortz und Döring [BD06] (Tabelle 9.1) in einem mittleren Bereich ($\eta^2 = 0.135$).

Eine post-hoc Testung (Bonferroni) zeigt signifikante Unterschiede nur für den Orientierungsschall 3 ($M = 1,98$ $SD = 0,96$) zum Orientierungsschall 1 ($M = 2,75$ $SD = 0,86$, $p < .001$), Orientierungsschall 2 ($M = 2,83$ $SD = 0,79$ $p < .001$) und Orientierungsschall 4 ($M = 2,58$ $SD = 0,94$ $p = .004$).

Analog zu den Überlegungen aus Abschnitt 4.5 wird die ANOVA ohne die Gruppe Experten 1 wiederholt. Der Faktor „Gruppe“ ist für diesen Fall zweistufig.

Es ergibt sich auch hier kein Einfluss ($F(1,120) < 1$) des Faktors „Gruppe“ auf die Daten. Der Effekt der Orientierungsschalle ($F(3,120) = 8.38$, $p < .001$, $\eta^2 = 0.173$) ist vergleichbar zu der Auswertung für alle drei Gruppen, ebenso wie die Interaktion der beiden Faktoren ($F(3,120) = 2.32$, $p = .078$).

Die post-hoc Testung (Bonferroni) deckt Unterschiede zwischen dem Orientierungsschall 3 ($M = 1,91$ $SD = 0,93$) und Orientierungsschall 1 ($M = 2,66$ $SD = 0,90$ $p = .006$) auf. Weiterhin unterscheidet sich Orientierungsschall 3 auch von Orientierungsschall 2 ($M = 3,00$ $SD = 0,76$ $p < .001$) und Orientierungsschall 4 ($M = 2,56$ $SD = 1,01$ $p = .024$).

Im Vergleich zu der nichtparametrischen Auswertung (Abschnitt 4.5) ergeben sich keine wesentlich neuen Erkenntnisse.

⁷² „Allgemeines lineares Modell, univariat, feste Faktoren, Quadratsummen vom Typ III“ [SPS02].

E. Anhang: THK von Einzeltönen in Rauschen im vollständigen Paarvergleich

E.1. Instruktion der Probanden im Wortlaut

Im Folgenden ist die Instruktion der Probanden, genauso wie sie für die Untersuchung benutzt wird, dokumentiert. Sie wird den Probanden zu Anfang einer Versuchsreihe am Bildschirm präsentiert, danach beginnt die Untersuchung:

Im Folgenden werden Ihnen 153 Geräuschpaare hintereinander vorgespielt, ein Paar besteht dabei jeweils aus zwei aufeinanderfolgenden Geräuschen.

Nachdem Sie ein solches Paar gehört haben besteht Ihre Aufgabe jeweils darin zu entscheiden, bei welchem von den beiden Geräuschen die TONHALTIGKEIT ausgeprägter ist.

Wenn Sie der Meinung sind, das 1. Geräusch war tonhaltiger, so drücken Sie die "1". War das 2. Geräusch tonhaltiger, dann drücken Sie bitte die "2".

Der Versuch wird erst dann fortgesetzt, wenn Sie Ihr Urteil abgegeben haben. Beachten Sie bitte, dass Ihnen jedes Paar nur einmal vorgespielt wird.

Um Ihnen eine Hilfe zu geben was für eine Hörempfindung der Begriff TONHALTIGKEIT beschreibt, wird Ihnen im Folgenden ein Beispiel für ein tonhaltiges Geräusch vorgespielt.

E.2. Berechnete Tonhaltigkeit der verwendeten Stimuli

Bei der Erstellung der Stimuli für die Paarvergleichsuntersuchung wird angestrebt, die in Tabelle 5.1 in der linken Spalte angegebenen maßgeblichen Dif-

ferenzen möglichst gut durch geeignete Kombination der Pegel von Tönen und Rauschen zu erreichen.

Tabelle E.1.: Berechnete THK (hier die maßgebliche Differenz ΔL in dB) für die 36 Stimuli. In der linken Spalte steht die bei der Erstellung der Stimuli beabsichtigte maßgebliche Differenz ΔL . In den Spalten daneben ist jeweils das nach DIN 45681:2006-08 berechnete ΔL eingetragen. Die Frequenz 50 Hz liegt außerhalb des Anwendungsbereiches der Norm (\clubsuit), für $\Delta L = \infty$ (Inf) dB ist die Norm nicht definiert (\checkmark) und für sehr große maßgebliche Differenzen ΔL funktioniert das Berechnungsverfahren nicht ohne Weiteres (\dagger), siehe Text.

	50 Hz	100 Hz	470 Hz	920 Hz	1450 Hz	2700 Hz
Inf	\clubsuit	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
48	\clubsuit	\dagger	\dagger	\dagger	\dagger	\dagger
36	\clubsuit	\dagger	32.6	34.8	34.3	35.0
24	\clubsuit	21.8	22.4	24.0	23.6	23.3
12	\clubsuit	11.8	11.3	11.2	11.3	11.3
0	\clubsuit	2.0	1.8	0	0	0

Da das Rauschen zeitlich fluktuiert, ist dieses Vorhaben nur mit einer gewissen Ungenauigkeit umzusetzen. Aus diesem Grund wird im Folgenden berechnet, welche maßgeblichen Differenzen die Versuchsstimuli tatsächlich aufweisen.

Um die Stimuli bzgl. ihrer maßgeblichen Differenz (d.h. der THK) zu analysieren, müssen die Stimuli zunächst verlängert werden, da sie mit einer Dauer von nur 500 ms nicht mit dem Verfahren nach DIN 45681:2006-08 [THK06] analysiert werden können. Die Verlängerung erfolgt, indem die Schalle dreimal hintereinander kopiert werden. Danach werden sie mit dem gleichen Verfahren wie in Abschnitt D.2 ausgewertet.

Für die Signale mit $\Delta L = \infty$ (Inf) dB macht es per Definition keinen Sinn, die Norm anzuwenden. Diese Signale sind daher in Tabelle E.1 mit \checkmark gekennzeichnet.

Für die Schalle mit einer Tonfrequenz von 50 Hz ist die Norm nicht geeignet, da sie erst ab 90 Hz angewendet werden kann. Diese Signale sind mit \clubsuit markiert.

Für sehr große maßgebliche Differenzen ΔL ist der Algorithmus für die Trennung von Ton und Rauschen mit den in der Norm definierten Parametern nicht ohne weiteres geeignet, siehe Kapitel 2.11. Die Stimuli können daher nicht analysiert werden und sind in Tabelle E.1 mit \dagger gekennzeichnet.

Die berechneten maßgeblichen Differenzen in Tabelle E.1 sind tendenziell etwas zu gering, bis auf die beiden Werte für $\Delta L = 0$ bei 100 Hz und 470 Hz. Für

diese beiden Frequenzen ist die Frequenzgruppenbreite relativ klein und daher fallen nur wenige Spektrallinien in diesen Bereich. Der Pegel des Rauschens in diesen Frequenzgruppen fluktuiert statistisch somit stärker als für höhere Frequenzgruppen. Dieser Effekt ist bei der Erstellung der Stimuli nicht kompensiert worden und trägt daher zu den festgestellten Abweichungen bei.

Die Unterschätzung der meisten anderen maßgeblichen Differenzen liegt zum einen daran, dass bei der Erstellung der Stimuli nicht der wirklich verwendete Abschnitt des Rauschens berücksichtigt worden ist, sondern die mittleren statistischen Werte des Rauschens für einen deutlich längeren Zeitraum.

Zum anderen wird der Schmalbandpegel des Rauschens in den Frequenzgruppen durch die DIN 45681:2006-08 mit steigendem ΔL überschätzt, so dass die THK dadurch unterschätzt wird.

F. Anhang: Einfluss verschiedener Parameter auf die THK

F.1. Verwendete Stimuli

Die Stimuli werden während des Experiments jedes Mal neu generiert. Dafür werden die internen Kommandos der Versuchssoftware SI [PK96] verwendet. Für das GAR wird jedes Mal ein neuer, eine Sekunde langer Abschnitt zufällig aus einem im Vorfeld der Studie generierten, längeren Signal herausgeschnitten und mit den tonalen Komponenten gemischt.

Die nachfolgenden Demonstrationen sind nach Abschluss der Studien mit Matlab [The06] den Versuchsstimuli nachgebildet worden. Sie sind nicht 100%ig identisch, entsprechen aber dem Höreindruck der verwendeten Stimuli. Das GAR für die folgenden Demonstrationen ist jeweils identisch.

Tabelle F.1.: Verwendete Stimuli.

Kap.	Versuch	Stimuli
A.1	– alle –	GAR
6.4.1	Anzahl Teiltöne „ohne Energieausgleich“	HKT(1), HKT(2), HKT(4), HKT(8)
6.4.2	Harmonizität Teiltöne	-300 Hz, -150 Hz, 0 Hz, 150 Hz, 300 Hz
6.4.3	Abstand Teiltöne	2 Hz, 10 Hz, 70 Hz, 140 Hz, 270 Hz
6.4.4	absoluter Darbietungspegel	HKT(1) in 7 Pegelstufen
6.4.5	Bandbreite SBR	2 Hz, 10 Hz, 70 Hz, 140 Hz, 270 Hz
6.4.6	Modulation (AM)	2 Hz, 10 Hz, 30 Hz, 70 Hz, 100 Hz
6.4.7	Modulation (FM)	2 Hz, 10 Hz, 30 Hz, 70 Hz, 100 Hz
6.4.8	Frequenzabhängigkeit	200 Hz, 350 Hz, 510 Hz, 700 Hz, 920 Hz, 1170 Hz, 1480 Hz, 1850 Hz, 2320 Hz, 2900 Hz, 3700 Hz
7.1	Anzahl Teiltöne „mit Energieausgleich“	HKT(1), HKT(2), HKT(4), HKT(8)

Durch Anklicken der Stimuli in Tabelle F.1 wird bei geeigneter Konfiguration und verfügbarem Zugriff auf das Internet der jeweilige Stimulus abgespielt. Alternativ können die Stimuli auch direkt von der Website www.tonhaltigkeit.de heruntergeladen oder angehört werden.

F.2. Verdeckungsmaß für die Teiltöne der HKTs

In den Kapiteln 6.4.1 und 7 werden HKTs mit unterschiedlich vielen Teiltönen (TT) in GAR verwendet. Während die DIN 45681E:1992-01 bei der Berechnung der maßgeblichen Differenz nach Gleichung 6.6 konstante Werte für alle Teiltöne ermittelt, ergeben sich für die DIN 45681:2006-08 nach Gleichung 6.4 frequenzabhängige Werte.

Daher ist in Tabelle F.2, neben der Frequenz für die TT, in der dritten Spalte das Verdeckungsmaß angegeben. In der vierten Spalte steht die maßgebliche Differenz für jeden der Einzeltöne bei Berechnung anhand der DIN 45681:2006-08 für die Stimuli in Kapitel 6.4.1.

In der letzten Spalte ist die Abweichung des Verdeckungsmaßes bzw. der maßgeblichen Differenz jedes Teiltönen zum Teilton mit der Nummer 1 angegeben.

Tabelle F.2.: Frequenz, Verdeckungsmaß, maßgebliche Differenz und Abweichung zum Grundton (siehe Text) für die acht Teiltöne der HKTs.

TT Nr.	Frequenz f_{TT} in Hz	Verdeckungsmaß $a_v(f_{TT})$ in dB	Maßgebl. Diff. ΔL_{TT} in dB	Abweichung in dB
1	700	-2,5	11,1	0
2	1400	-3,2	11,8	0,7
3	2100	-3,6	12,2	1,1
4	2800	-3,9	12,5	1,4
5	3500	-4,1	12,7	1,6
6	4200	-4,3	12,9	1,8
7	4900	-4,5	13,1	2,0
8	5600	-4,6	13,2	2,1

F.3. Erweiterte Datenanalyse (ANOVA)

Unter der Annahme, dass die Ergebnisse des adaptiven THK-Messverfahrens Intervallskalenniveau besitzen, kann mit den Daten aus Kapitel 6.3 post-hoc für jeden der untersuchten Parameter eine zweifaktorielle ANOVA (messwiederholt) durchgeführt werden.

Ein Faktor (vier- bis elfstufig) beschreibt jeweils den Effekt des untersuchten Parameters auf die THK, der andere Faktor (zweistufig) gibt den Effekt des Startpegels wieder.

Kann kein Effekt des Startpegels nachgewiesen werden, so ist es gerechtfertigt, die Daten der beiden Startpegel zusammenzufassen und gemeinsam darzustellen und auszuwerten. Das geschieht, indem die individuellen Daten pro Proband über die beiden Startpegel zunächst gemittelt werden, bevor die nächsten Analyseschritte erfolgen.

Die individuellen Probandendaten werden anschließend zu Mittelwerten mit Standardabweichung zusammengefasst und in allen folgenden Abbildungen jeweils über den untersuchten Parameter aufgetragen. Die THK-Daten sind aus Gründen der Vergleichbarkeit genau so dargestellt, wie bereits in Abschnitt 6.3.3 ausführlich beschrieben.

Für den Fall, dass die ANOVA einen Effekt des untersuchten Parameters offenbart, werden paarweise T -Tests für gepaarte Stichproben (für die gemittelten Startpegel) durchgeführt. Die Paare, die sich ohne Korrektur der Inflation des α -Fehlers⁷³ signifikant voneinander unterscheiden, sind jeweils tabellarisch aufgeführt.

Diese Tabellen für die T -Tests enthalten jeweils von links nach rechts das Stimulus-Paar, den T -Wert, die Freiheitsgrade (df), die unkorrigierte Signifikanz und die Signifikanz bei Korrektur des α -Fehlers durch die Bonferroni-Methode.

Ergänzend wird die Effektgröße⁷⁴ η^2 berechnet. Sie beschreibt die *Varianzaufklärung*, d.h. „denjenigen Varianzanteil der abhängigen Variable, der auf die unabhängige Variable zurückzuführen ist“ ([Bor93], Seite 257). Die Effektgröße nimmt einen Wert aus dem Intervall [0:1] an.

Die Effektgröße macht in erster Linie Aussagen über Unterschiede von Mittelwerten. Im Gegensatz dazu beziehen sich die Aussagen bei Betrachtung der „Signifikanz“ eher auf die Breite der vorliegenden Verteilungen. Die Effektgröße sollte sich daher nicht oder nur wenig verändern, wenn der Stichprobenumfang vergrößert wird [RR99]. Im Gegensatz dazu sind die Aussagen bzgl. Signifikanzen und Signifikanzniveaus erheblich vom Stichprobenumfang abhängig.

In diesem Zusammenhang sei nochmal darauf hingewiesen, dass bei allen Arten von statistischen Tests neben der reinen Betrachtung von Effektgrößen oder Signifikanzen immer die praktische Bedeutsamkeit berücksichtigt werden muss [BD06]. Darauf wird in Kapitel 6.4 besonderer Wert gelegt.

Für die Effektgröße bei Fragestellungen zur Varianzaufklärung gilt nach Bortz und Döring [BD06] (Tabelle 9.1) die folgende Klassifikation der Effektgrößen: $\eta^2 < 0.01$ beschreibt einen kleinen Effekt, $\eta^2 > 0.25$ einen großen Effekt und Werte

⁷³ Werden Paare verglichen, ohne die α -Fehler Inflation zu berücksichtigen, wird dies im Folgenden zumeist als „unkorrigierte Signifikanz“ bezeichnet.

⁷⁴ Die Effektgröße wird synonym auch als *Effektstärke* bezeichnet.

dazwischen einen mittleren Effekt.

Im Folgenden werden die so gewonnenen Ergebnisse grafisch dargestellt, ohne im einzelnen eine ausführliche Bewertung oder Diskussion vorzunehmen. Eine kritische Einschätzung wird für die Ergebnisse, für die es sinnvoll erscheint, in Abschnitt 6.4 durchgeführt.

F.3.1. Anzahl von Tönen

Ein Effekt des Startpegels auf die Bewertung der THK kann nicht nachgewiesen werden ($F(1,4)<1$), daher werden die beiden Startpegel durch Mittelung zusammengefasst. Den Mittelwert über alle Probanden samt Standardabweichung zeigt Abbildung F.1.

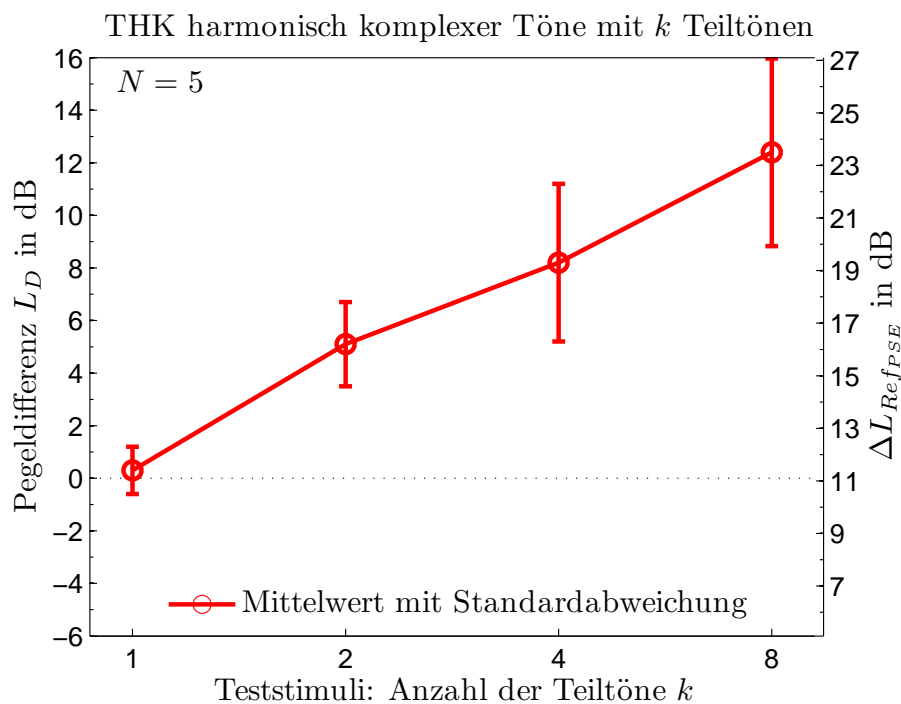


Abbildung F.1.: THK harmonisch komplexer Töne (HKT) ohne Energieausgleich für fünf Probanden, aufgetragen über die Anzahl der Teiltöne k des Teststimulus.

Für den vorliegenden Datensatz zeigt sich eine signifikante Abhängigkeit der Daten von der Anzahl der Teiltöne k der HKTs ($F(3,12)=25.1$, $p<.001$).

Die post-hoc paarweise durchgeführten T-Tests für gepaarte Stichproben (Mittelwert der beiden Startpegel) sind in Tabelle F.3 zusammengefasst.

Ohne Korrektur des α -Fehlers unterscheiden sich alle Stimuli signifikant voneinander ($p<.049$), siehe vierte Spalte von Tabelle F.3.

Tabelle F.3.: T-Tests für die HKT-Stimuli ohne Energieausgleich.

Paar	T-Wert	<i>df</i>	Signifikanz (2-seitig)	Signifikanz, Bonferroni (2-seitig)
$k=1 ; k=2$	-5,16	4	.007	.040
$k=1 ; k=4$	-5,09	4	.007	.042
$k=1 ; k=8$	-7,38	4	.002	.011
$k=2 ; k=4$	-3,39	4	.027	.165
$k=2 ; k=8$	-3,98	4	.016	.098
$k=4 ; k=8$	-2,79	4	.049	.296

Bei Berücksichtigung des α -Fehlers (Bonferroni) ergeben sich signifikante Unterschiede auf dem 5%-Niveau nur für den ersten Stimuli ($k=1$) zu den anderen drei Stimuli. Alle anderen Stimulus-Paare unterscheiden sich nicht.

Die Effektgröße für die vorliegenden Daten ist hoch: $\eta^2=.863$. D.h. mehr als 86% der Varianz der vorliegenden Daten wird durch die Anzahl der Teiltöne k aufgeklärt.

F.3.2. Harmonizität von Teiltönen

Ein Effekt des Startpegels auf die Daten ist nicht nachzuweisen ($F(1,4)<1$), daher sind die Mittelwerte samt Standardabweichungen für die Daten in Abbildung F.2 dargestellt.

Es zeigt sich keine signifikante Abhängigkeit der THK von der Frequenzverschiebung $f_{shift}(F(4,16)=1.8, p=.19)$, T-Tests werden daher nicht durchgeführt.

F.3.3. Abstand zweier Sinustöne

Ein Effekt des Startpegels auf die THK kann mit $F(1,6)<1$ nicht nachgewiesen werden. Aus diesem Grund sind die Daten für die beiden Startpegel durch Mittelung zusammengefasst und in Abbildung F.3 dargestellt.

Der Abstand zweier Sinustöne hat einen signifikanten Effekt auf die THK ($F(4,24)=14.6, p<.001$). Post-hoc T-Tests zwischen den fünf Stimuli ergeben die in Tabelle F.4 aufgelisteten Werte.

Der Frequenzabstand von 2 Hz unterscheidet sich nach Bonferroni-Korrektur signifikant von den Frequenzabständen 140 Hz und 270 Hz, 10 Hz unterscheidet sich von 140 Hz.

Die Effektgröße als Maß für die Varianzaufklärung ist mit $\eta^2=.708$ auch hier sehr groß.

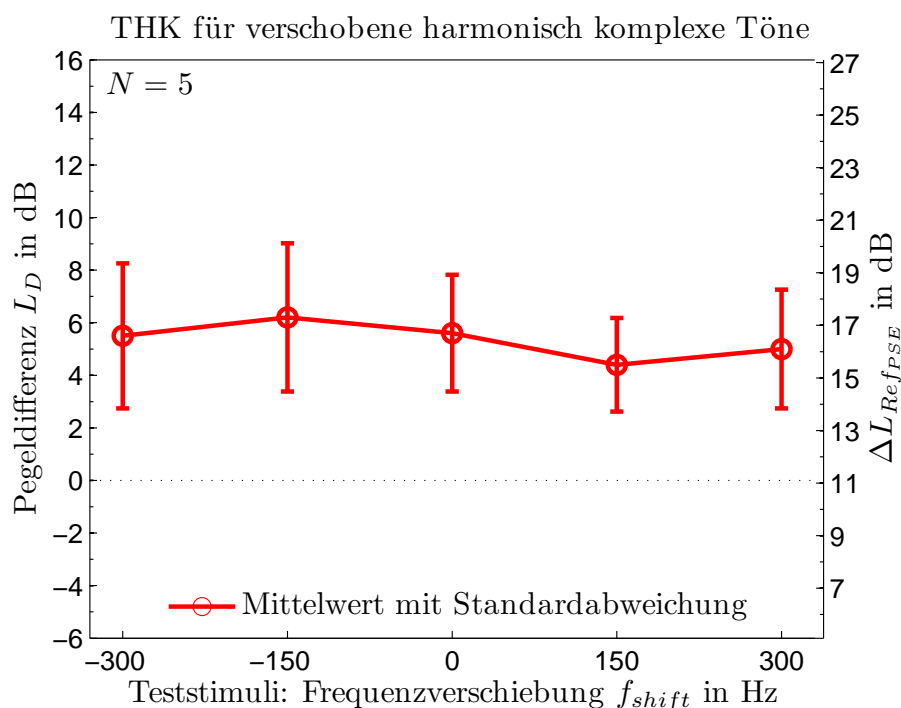


Abbildung F.2.: THK verschobener harmonisch komplexer Töne (VHKT) für fünf Probanden, aufgetragen über die Frequenzverschiebung f_{shift} des Teststimulus.

Tabelle F.4.: Relevante T-Tests für die Abstände der Sinustöne.

Paar	T-Wert	df	Signifikanz (2-seitig)	Signifikanz, Bonferroni (2-seitig)
2; 70	3.27	6	.017	.171
2; 140	6.02	6	.001	.009
2; 270	5.07	6	.002	.023
10; 140	5.86	6	.001	.011
10; 270	4.14	6	.006	.061
70; 140	3.10	6	.021	.210
70; 270	2.87	6	.028	.285
140; 270	2.87	6	.028	.285

F.3.4. Absoluter Darbietungspegel

Es liegt kein Effekt des Startpegels auf die THK vor ($F(1,7)=1.29$ $p=.294$). Die gemittelten Daten sind in Abbildung F.4 dargestellt.

Für die THK zeigt sich eine signifikante Abhängigkeit vom absoluten Darbie-

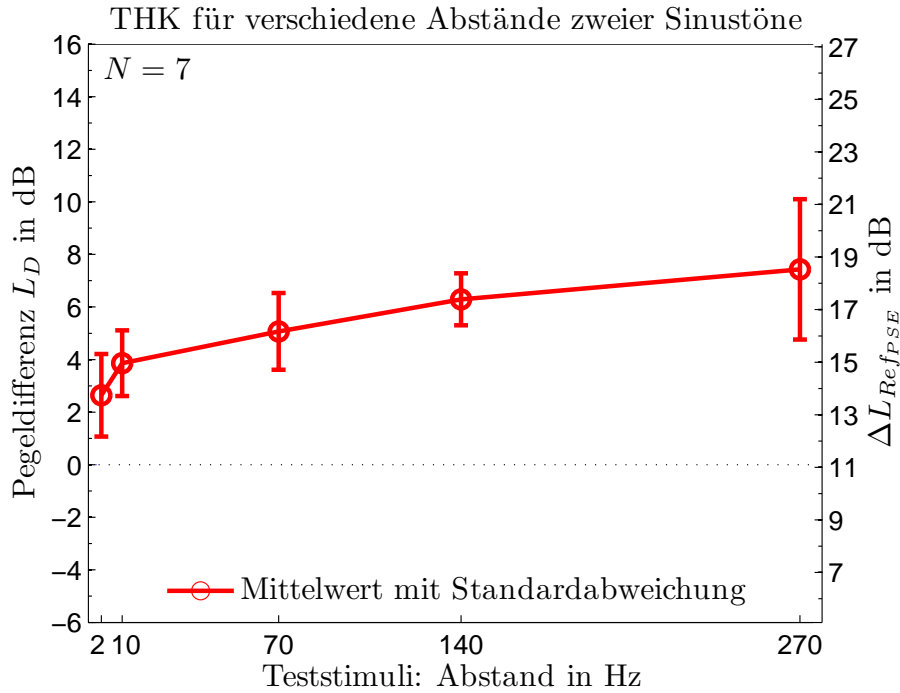


Abbildung F.3.: Effekt des Abstandes von zwei Sinustönen auf die THK für sieben Probanden. Aufgetragen ist Mittelwert und Standardabweichung über den Abstand der Sinustöne des Teststimulus.

tungspegel ($F(6,42)=5.12, p<.001$). Die Effektgröße ergibt sich hier zu $\eta^2=.423$. Post-hoc durchgeführte T-Tests sind in Tabelle F.5 zusammengefasst.

Tabelle F.5.: Relevante T-Tests für die sieben Darbietungspegel.

Paar	T-Wert	df	Signifikanz (2-seitig)	Signifikanz, Bonferroni (2-seitig)
-30 ; -20	-3.20	7	.015	.315
-30 ; 0	-5.18	7	.001	.027
-30 ; 10	-6.66	7	.000	.006
-30 ; 20	-4.19	7	.004	.086
-20 ; 10	-2.73	7	.029	.618
-10 ; 0	-3.21	7	.015	.312
-10 ; 10	-3.03	7	.019	.399
-10 ; 20	-2.59	7	.036	.752

Die beiden Darbietungspegel -30 dB und -10 dB scheinen im Wesentlichen den Effekt des absoluten Darbietungspegels auf die THK hervorzurufen, siehe auch Diskussion in Abschnitt 6.4.4.

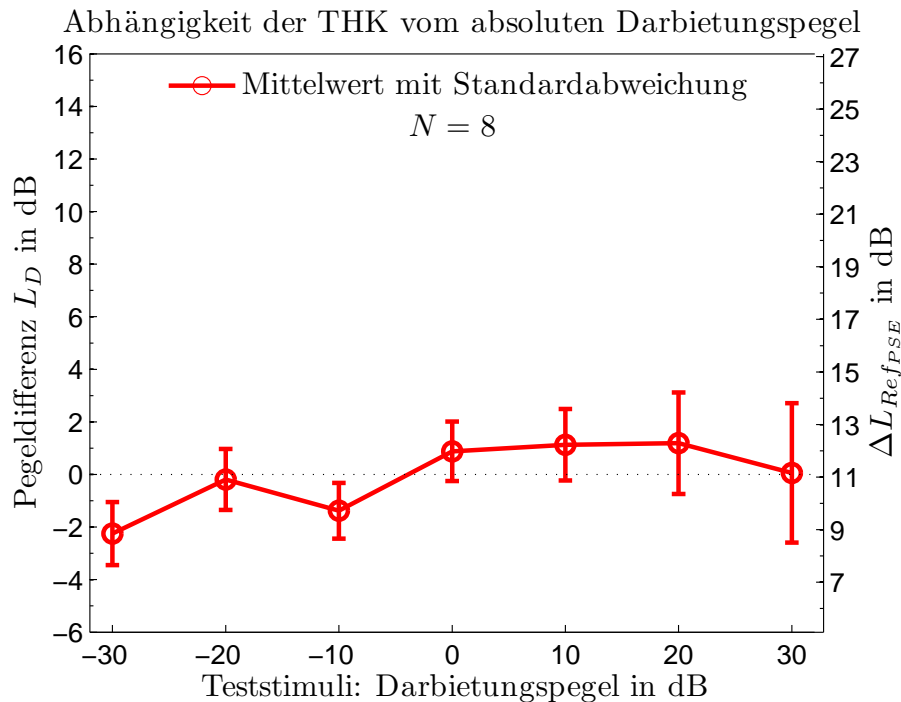


Abbildung F.4.: Mittelwert und Standardabweichung der THK von acht Probanden, aufgetragen über den absoluten Darbietungspegel des Teststimulus.

Eine Wiederholung der Analyse ohne den Darbietungspegel von -30 dB ergibt unverändert keine Abhängigkeit der THK vom Startpegel ($F(1,7) < 1$). Der Effekt des absoluten Darbietungspegels auf die THK fällt jedoch deutlich geringer aus ($F(5,35) = 2.83$, $p = .030$), ebenso die Effektgröße $\eta^2 = .288$.

Eine Analyse ohne die Darbietungspegel von -30 dB und -10 dB ergibt keine Abhängigkeit der THK vom Startpegel ($F(1,7) = 2.17$, $p = .184$) oder vom absoluten Darbietungspegel ($F(4,28) = 1.09$, $p = 0.383$).

F.3.5. Bandbreite von Schmalbandrauschen

Es existiert kein Effekt des Startpegels auf die THK ($F(1,6) < 1$), die Daten mit Mittelwert und Standardabweichung zeigt Abbildung F.5.

Für den vorliegenden Datensatz zeigt sich eine signifikante Abhängigkeit der THK von der Breite des Schmalbandrauschens ($F(4,24) = 10.10$, $p < .001$). Post-hoc durchgeführte T-Tests sind in Tabelle F.6 zusammengefasst. Die Effektgröße beträgt $\eta^2 = .627$.

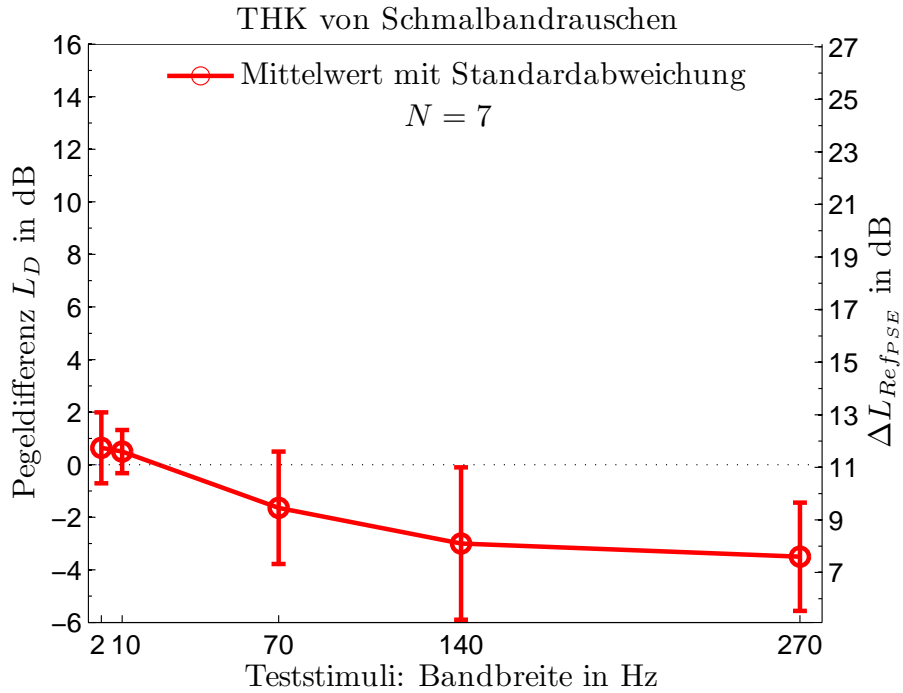


Abbildung F.5.: THK für sieben Probanden über die Breite des Schmalbandrauschens (Teststimulus).

Tabelle F.6.: Relevante T-Tests für die Breite des Schmalbandrauschens.

Paar	T-Wert	df	Signifikanz (2-seitig)	Signifikanz, Bonferroni (2-seitig)
2; 70	-2.75	6	.033	.331
2; 140	-3.72	6	.010	.099
2; 270	-4.36	6	.005	.048
10; 70	-3.12	6	.021	.207
10; 140	-3.34	6	.016	.155
10; 270	-5.59	6	.001	.014
70; 270	-3.23	6	.018	.178

F.3.6. Modulationsfrequenz (AM) modulierter Töne

Ein Effekt des Startpegels auf die THK ($F(1,6) < 1$) kann nicht nachgewiesen werden. Die gemittelten Daten zeigt Abbildung F.6.

Für den vorliegenden Datensatz zeigt sich eine signifikante Abhängigkeit der THK von der Modulationsfrequenz ($F(4,24) = 3.24$, $p = .029$). Post-hoc durchgeführte T-Tests sind in Tabelle F.7 aufgelistet.

Die Effektgröße $\eta^2 = .351$ ist gemessen an den anderen Untersuchungen ver-

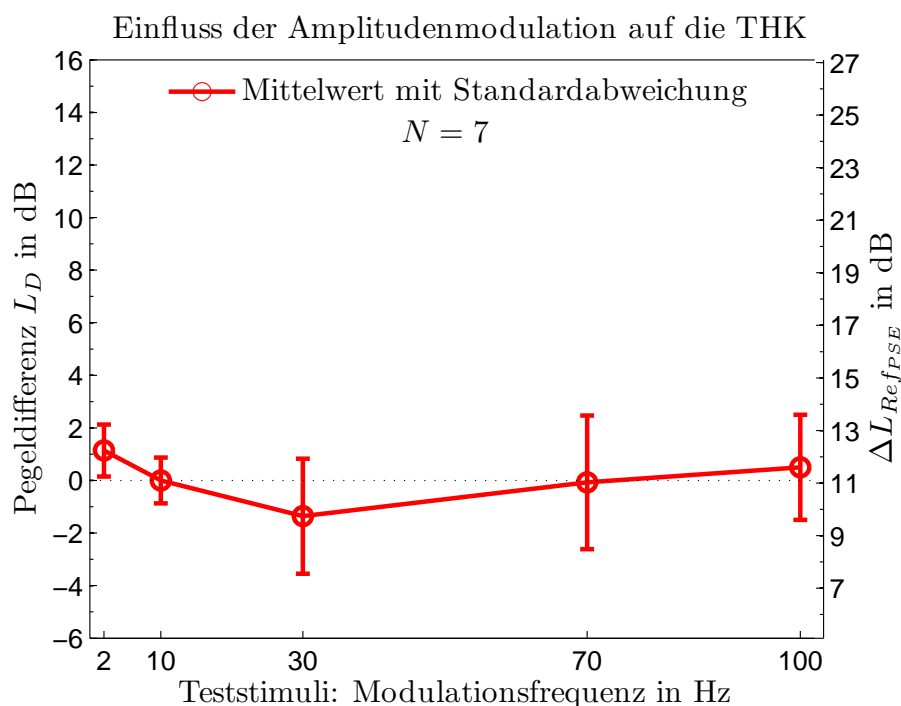


Abbildung F.6.: THK amplitudenmodulierter Töne für sieben Probanden, aufgetragen über die Modulationsfrequenz des Teststimulus.

Tabelle F.7.: Relevante T-Tests für amplitudenmodulierte Töne (Modulationsfrequenz).

Paar	T-Wert	df	Signifikanz (2-seitig)	Signifikanz, Bonferroni (2-seitig)
2; 10	2.73	6	.034	.342
2; 30	2.93	6	.026	.262
30; 70	-2.46	6	.049	.488
30; 100	-3.36	6	.015	.153

gleichsweise gering, aber nach Bortz und Döring [BD06] immer noch als „groß“ einzustufen.

F.3.7. Modulationsfrequenz (FM) modulierter Töne

Ein Effekt des Startpegels auf die THK ($F(1,6) < 1$) kann nicht nachgewiesen werden. Die Mittelwerte samt Standardabweichung der Daten zeigt Abbildung F.7.

Für den vorliegenden Datensatz zeigt sich eine signifikante Abhängigkeit der

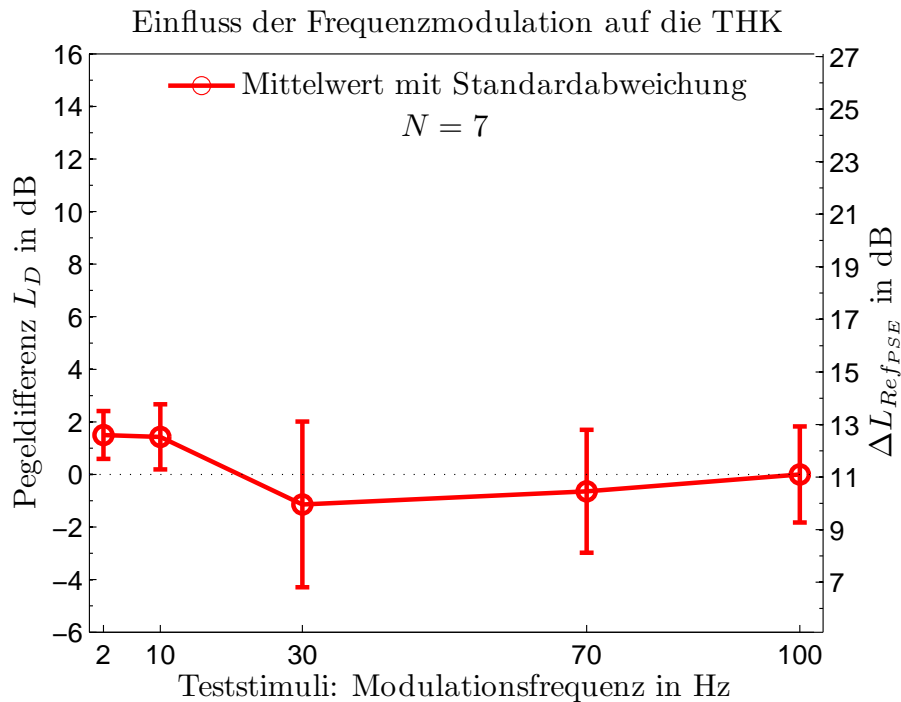


Abbildung F.7.: THK frequenzmodulierter Töne für sieben Probanden, aufgetragen über der Modulationsfrequenz des Teststimulus.

THK von der Modulationsfrequenz ($F(4,24)=4.17, p=.010$). Paarweise T-Tests für die fünf Stimuli sind in Tabelle F.8 aufgelistet. Die Effektgröße beträgt $\eta^2=.410$.

Tabelle F.8.: Relevante T-Tests für frequenzmodulierte Töne (Modulationsfrequenz).

Paar	T-Wert	df	Signifikanz (2-seitig)	Signifikanz, Bonferroni (2-seitig)
10; 70	3.10	6	.021	.210
10; 100	2.65	6	.038	.381

F.3.8. Tonfrequenz

Diese Analysen werden für zwei verschiedene Vergleichsstimuli durchgeführt, vgl. Abschnitt 6.4.8 und 6.4.9.

Sowohl für den Vergleichsstimulus mit 700 Hz als auch für den Stimulus mit 1850 Hz existiert kein Effekt des Startpegels auf die THK, in beiden Fällen gilt

F. Einfluss verschiedener Parameter auf die THK

$F(1,7) < 1$. Die gemittelten Daten für beide Fälle zeigt Abbildung F.8 analog zu Abbildung 6.13 aus Abschnitt 6.4.9.

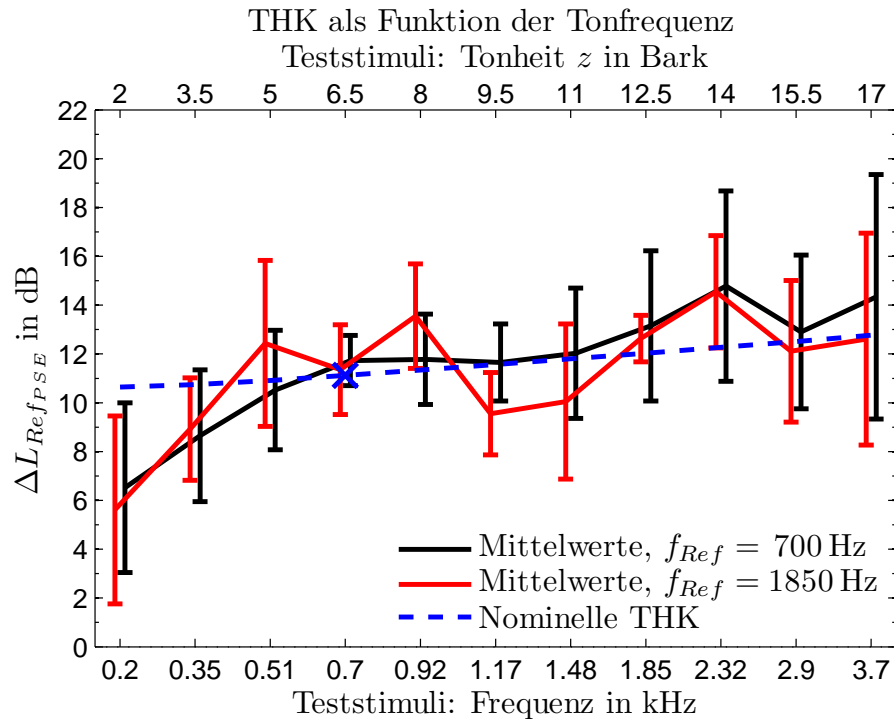


Abbildung F.8.: THK für acht Probanden, aufgetragen über die Tonfrequenz des Teststimulus.

Für beide Vergleichsstimuli zeigt sich eine signifikante Abhängigkeit der THK von der Frequenz. Für den Vergleichsstimulus mit 700 Hz gilt $F(10,70)=6.32$, $p < .001$ und für den Stimulus mit 1850 Hz $F(10,70)=7.02$, $p < .001$.

Post-hoc durchgeführte T-Tests sind für den Vergleichsstimulus mit 700 Hz in Tabelle F.9 aufgeführt, für den Vergleichsstimulus mit 1850 Hz in Tabelle F.10.

Die Effektgröße liegt für beide Vergleichsstimuli in der gleichen Größenordnung. Für den Vergleichsstimulus mit 700 Hz beträgt sie $\eta^2 = .474$, für den Vergleichsstimulus mit 1850 Hz ist die Varianzaufklärung etwas größer: $\eta^2 = .501$.

Tabelle F.9.: Relevante T-Tests für die Tonfrequenzen (Vergleichsstimulus 7000 Hz).

Paar	T-Wert	df	Signifikanz (2-seitig)	Signifikanz, Bonferroni (2-seitig)
200 ; 510	-3.91	7	.006	.320
200 ; 700	-4.32	7	.003	.192
200 ; 920	-3.22	7	.015	.801
200 ; 1170	-4.61	7	.002	.134
200 ; 1480	-4.58	7	.003	.139
200 ; 1850	-6.07	7	.001	.028
200 ; 2320	-4.73	7	.002	.117
200 ; 2900	-4.76	7	.002	.113
200 ; 3700	-3.96	7	.005	.299
350 ; 510	-7.85	7	.000	.006
350 ; 700	-4.53	7	.003	.148
350 ; 920	-4.09	7	.005	.254
350 ; 1170	-3.11	7	.017	.934
350 ; 1480	-6.37	7	.000	.021
350 ; 1850	-5.52	7	.001	.049
350 ; 2320	-5.25	7	.001	.065
350 ; 2900	-4.35	7	.003	.185
350 ; 3700	-3.71	7	.008	.416
510 ; 2320	-2.73	7	.029	1.00
700 ; 2320	-2.76	7	.028	1.00
920 ; 2320	-3.49	7	.010	.554
1170 ; 2320	-3.49	7	.010	.555
1480 ; 2320	-3.59	7	.009	.485
2320 ; 2900	2.39	7	.048	1.00

F. Einfluss verschiedener Parameter auf die THK

Tabelle F.10.: Relevante T-Tests für die Tonfrequenzen (Vergleichsstimulus 1850 Hz).

Paar	T-Wert	<i>df</i>	Signifikanz (2-seitig)	Signifikanz, Bonferroni (2-seitig)
200 ; 350	-4.25	7	.004	.208
200 ; 510	-4.48	7	.003	.158
200 ; 700	-5.75	7	.001	.038
200 ; 920	-6.51	7	.000	.018
200 ; 1170	-3.50	7	.010	.547
200 ; 1480	-3.32	7	.013	.705
200 ; 1850	-4.77	7	.002	.112
200 ; 2320	-4.97	7	.002	.089
200 ; 2900	-3.10	7	.017	.947
200 ; 3700	-2.98	7	.021	1.00
350 ; 510	-2.47	7	.043	1.00
350 ; 700	-3.67	7	.008	.439
350 ; 920	-5.48	7	.001	.051
350 ; 1850	-4.38	7	.003	.178
350 ; 2320	-4.48	7	.003	.158
700 ; 920	-3.07	7	.018	.996
700 ; 1170	3.70	7	.008	.423
700 ; 2320	-2.97	7	.021	1.00
920 ; 1170	4.56	7	.003	.142
920 ; 1480	2.65	7	.033	1.00
1170 ; 1850	-3.79	7	.007	.374
1170 ; 2320	-4.59	7	.003	.139
1480 ; 2320	-3.60	7	.009	.481
1850 ; 2320	-2.68	7	.032	1.00
2320 ; 2900	2.54	7	.039	1.00

G. Anhang: Verfahrensvergleich

G.1. Normierung der verschiedenen Vergleichsstimuli anhand der individuellen Verdeckungsmaße

Es folgen zwei Abbildungen (G.1 und G.2), die die Normierung der vier verschiedenen Vergleichsstimuli $HKT(k)$ anhand der individuellen Verdeckungsmaße $a_v(k)$ verdeutlichen. Für Erläuterungen dazu siehe Kapitel 7.6.2.

G.2. Transformation der Messwerte für Blocked Tracks

Die 16 einzelnen Messwerte für die THK mit Blocked Tracks sind in Abbildung G.3 über die Teststimuli (TS) aufgetragen. Die vier Vergleichsstimuli (VS) werden in diesem Fall jeweils durch einen Linienzug dargestellt.

Die Zahlenwerte der THK (L_D) aus Abbildung G.3 sind in Tabelle G.1 als Mittelwerte über die elf Probanden samt Standardabweichung tabelliert.

Tabelle G.1.: Mittelwerte und Standardabweichung der THK (L_D) für die 16 Messungen des Messdesign BT. Die Teststimuli (TS) sind in den Zeilen, die Vergleichsstimuli (VS) in den Spalten aufgetragen. Die Zahl hinter der Bezeichnung der Stimuli gibt die Anzahl der Teiltöne der $HKT(k)$ an. Alle Angaben in dB.

	VS 1	VS 2	VS 4	VS 8
TS 1	0,46 ±1,78	-0,55 ±2,27	-1,94 ±2,92	-2,92 ±3,05
TS 2	2,44 ±3,69	0,07 ±1,92	-1,04 ±1,69	-2,12 ±2,71
TS 4	2,30 ±4,37	0,55 ±2,32	-0,19 ±2,47	-1,05 ±1,72
TS 8	3,22 ±5,39	1,66 ±3,66	0,77 ±1,80	0,41 ±1,93
S	-2,57 ±1,64	-0,87 ±1,05	0,18 ±0,86	0,58 ±0,68
S ₀	0	1,70 ±1,05	2,75 ±0,86	3,15 ±0,68

Die grau hinterlegten Zellen (Diagonale der Matrix) sollten bei einer idealen

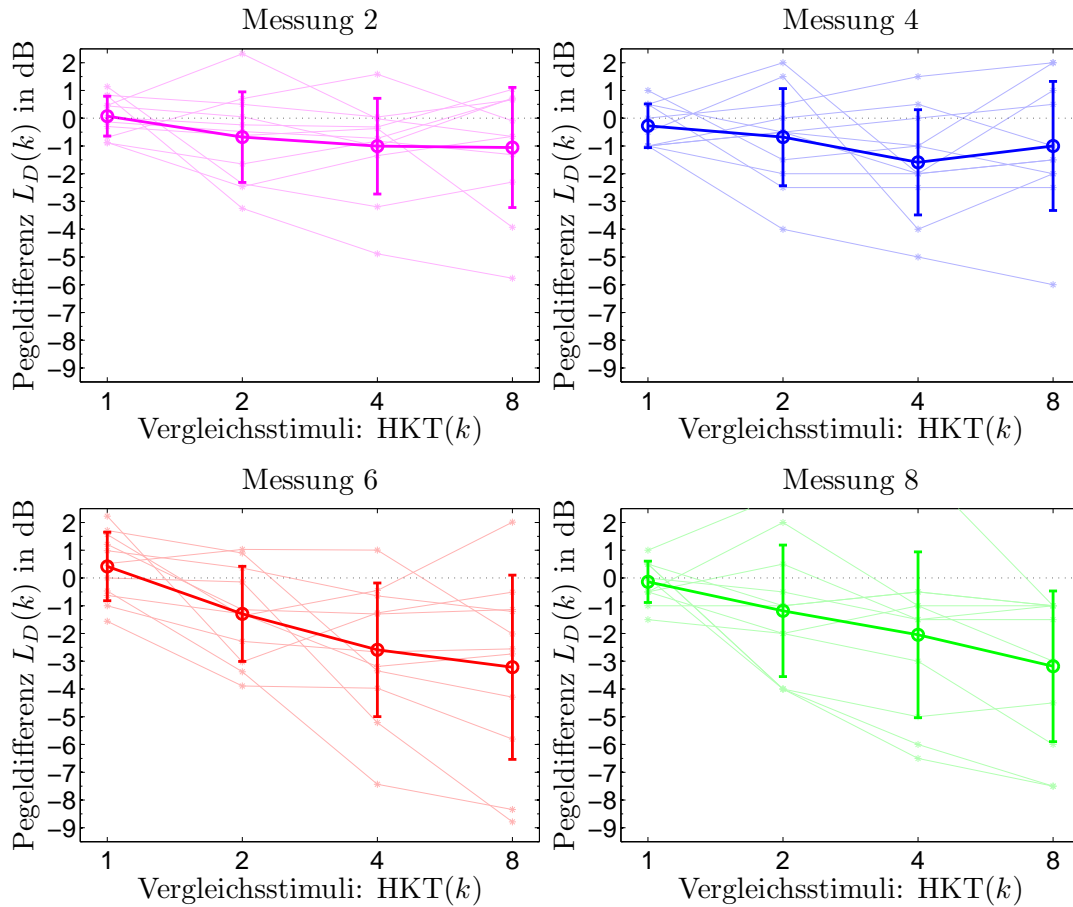


Abbildung G.1.: Gemessene Pegeldifferenzen $L_D(k)$ für Messungen 2, 4, 6 und 8. Für diese Messungen wird das THK-Referenzsignal als Teststimulus verwendet und anhand der zu untersuchenden Signale (als Vergleichsstimuli) skaliert. Die dünnen Linien geben den individuellen Verlauf für die elf Probanden wieder, die dicken Linien repräsentieren die Mittelwerte samt Standardabweichung.

Messung Null ergeben. Das ist hier in guter Näherung erfüllt.

In der mit „S” bezeichneten Zeile sind die Mittelwerte der Mithörschwelle für die $HKT(k)$ samt Standardabweichung verzeichnet, so wie in Kapitel 7.4 gemessen und diskutiert. In der Zeile „S₀” sind diese Werte auf die Mithörschwelle von $HKT(1)$ (Spalte V1) normiert.

Die Mittelwerte der Mithörschwelle (Zeile „S₀”) werden nun spaltenweise zu den Zellen der Tabelle G.1 addiert, die Standardabweichungen werden zusammengefasst (Gaußsche Fehlerfortpflanzung). Man erhält so das in Tabelle G.2 dargestellte Ergebnis.

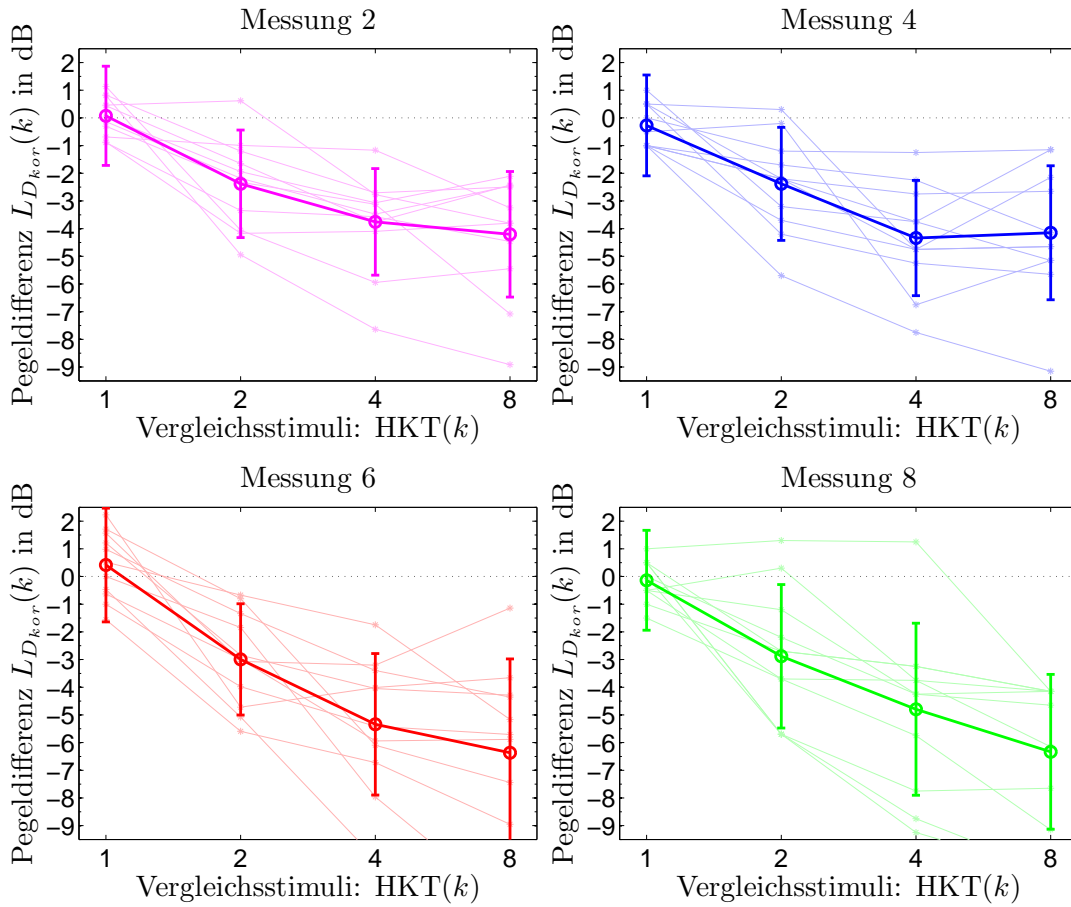


Abbildung G.2.: Korrigierte Pegeldifferenzen $L_{D_{kor}}(k)$ für Messungen 2, 4, 6 und 8. Die Korrektur berücksichtigt die individuellen Verdeckungsmaße $a_v(k)$ für die HKTs.

Die angegebene Standardabweichung berücksichtigt die zusammengefassten Standardabweichungen für die Messung der Schwelle und der THK (Gauß'sche Fehlerfortpflanzung).

Diese Daten sind in Abbildung G.4 grafisch dargestellt, die Skalierung ist nach wie vor auf L_D bezogen.

Die Spaltenmittelwerte (MW_S) der Tabelle G.2 sollten nach dieser Transformation möglichst nah beieinander liegen. Die ersten drei Spalten unterscheiden sich fast nicht, die vierte Spalte ist etwas kleiner als die anderen drei.

Die Zeilenmittelwerte (MW_Z) geben das Endergebnis der THK-Messung mit Blocked Tracks wieder, alle Angaben in dB.

In Tabelle G.3 sind diese THK-Werte zu absoluten THK-Angaben (bezogen auf ΔL_{RefPSE}) umgerechnet.

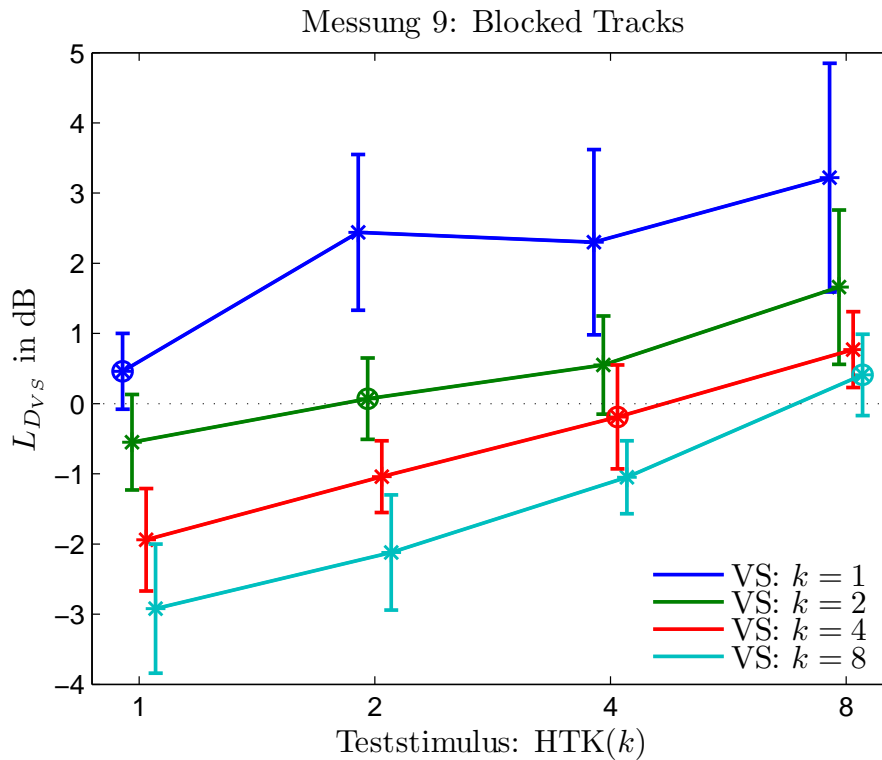


Abbildung G.3.: Rohdaten der THK (L_D) für die Messung mit Blocked Tracks über die Teststimuli (TS). Aufgetragen sind Mittelwerte und Standardfehler über elf Probanden. Durch einen Kreis sind jeweils die Konstellationen markiert, in denen Vergleichsstimulus (VS) und Teststimulus identisch sind.

Tabelle G.2.: Transformierte Messwerte für die Blocked Tracks. Die Messwerte sind hier spaltenweise um den Wert der gemessenen Mithörschwelle korrigiert. Teststimuli (TS) sind zeilenweise, Vergleichsstimuli (VS) spaltenweise aufgetragen. Alle Angaben in dB.

	VS1	VS2	VS4	VS8	(MW _Z)
TS 1	0,46 ± 2,42	1,15 ± 2,50	0,81 ± 3,04	0,23 ± 3,12	0,66 ± 2,79
TS 2	2,44 ± 4,04	1,77 ± 2,19	1,71 ± 1,90	1,03 ± 2,79	1,74 ± 2,85
TS 4	2,30 ± 4,67	2,25 ± 2,55	2,56 ± 2,62	2,10 ± 1,85	2,30 ± 3,10
TS 8	3,22 ± 5,64	3,36 ± 3,81	3,52 ± 1,99	3,56 ± 2,05	3,41 ± 3,69
(MW _S)	2,11 ± 4,35	2,13 ± 2,83	2,15 ± 2,43	1,73 ± 2,51	2,03 ± 3,13

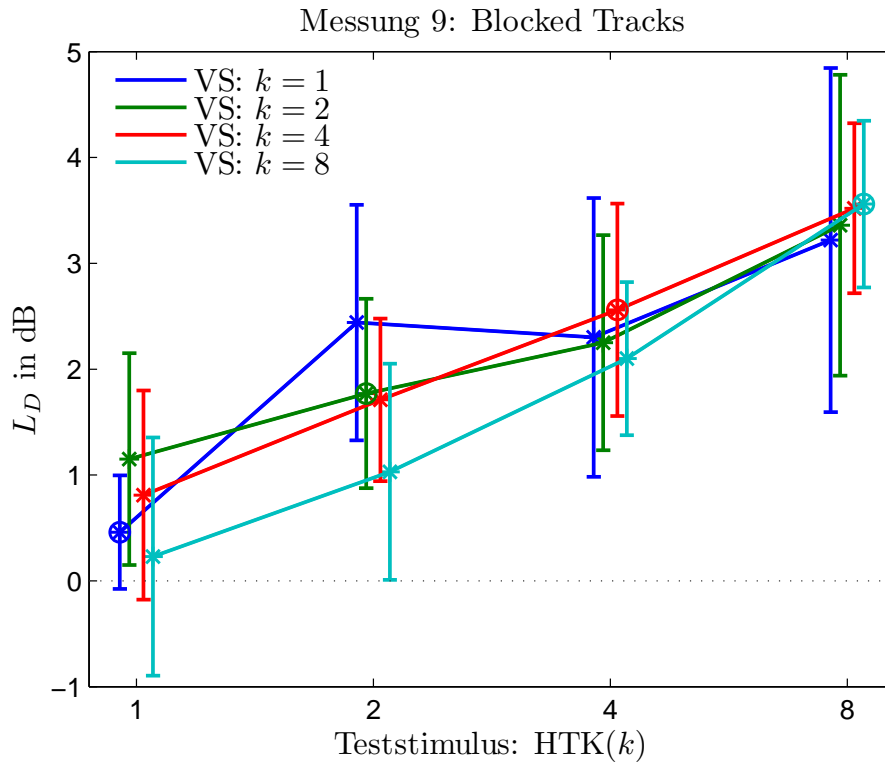


Abbildung G.4.: Transformierte THK-Daten für Blocked Tracks. Die Daten der Abbildung G.3 werden mittels der individuellen Verdeckungsmaße der $HKT(k)$ auf den $HKT(1)$ normiert. Aufgetragen sind Mittelwerte über die elf Probanden samt Standardfehler. Die Standardfehler berücksichtigen sowohl den Anteil der Verdeckungsmaße wie auch den Anteil der THK (Gauß'sche Fehlerfortpflanzung).

Tabelle G.3.: Absolute Messwerte für die Blocked Tracks (Mittelwerte und Standardabweichung der THK). Aufgetragen ist die THK (ΔL_{RefPSE}) jeweils zeilenweise für die vier Teststimuli, d.h. die $HKT(k)$. Alle Angaben in dB.

	VS1	VS2	VS4	VS8
TS 1	16,76 ± 2,42	17,45 ± 2,50	17,11 ± 3,04	16,53 ± 3,12
TS 2	18,74 ± 4,04	18,07 ± 2,19	18,01 ± 1,90	17,33 ± 2,79
TS 4	18,60 ± 4,67	18,55 ± 2,55	18,86 ± 2,62	18,40 ± 1,85
TS 8	19,52 ± 5,64	19,66 ± 3,81	19,82 ± 1,99	19,86 ± 2,05

Abbildungsverzeichnis

3.1. Individuelle Wahrnehmungsräume der THK	25
3.2. Gemittelter Wahrnehmungsraum der THK	26
3.3. Scree-Plot für den gemittelten Wahrnehmungsraum	26
3.4. Individuelle Rangreihen der Stimuli bzgl. ausgewählter Kontrastpaare	28
3.5. Skalierung der Stimuli anhand ausgewählter Kontrastpaare	30
4.1. Zeitverläufe der potentiellen Orientierungsschalle	35
4.2. Kategorialskala zur Bewertung der Orientierungsschalle	37
4.3. Häufigkeit der verwendeten Kategorien	38
4.4. Häufigkeit der verwendeten Kategorien über Schalle/Gruppen	39
4.5. Häufigkeit der Ränge für die Probandenurteile	40
4.6. Auf Intervallskalenniveau transformierte Probandenurteile	41
4.7. Standard-Orientierung: Grundlegende Schritte	45
4.8. Standard-Orientierung: Definition der THK	45
4.9. Standard-Orientierung: Darbietung des Orientierungsschalls	46
4.10. Standard-Orientierung: Ergänzende Orientierungsschalle	47
4.11. Standard-Orientierung: Abschließende Nachfrage	47
5.1. Übersicht der individuellen THK-Daten über ΔL	53
5.2. Intervallskalierte THK für Gruppe 1 und 2 individuell	55
5.3. Intervallskalierte THK für Gruppe 1 und 2 zusammengefasst	56
5.4. Übersicht der individuellen THK-Daten über f	58
5.5. Intervallskalierte THK für Gruppe 3 und 4	59
5.6. Intervallskalierte THK für Gruppe 3 und 4 ohne Inf	60
5.7. Intervallskalierte THK für Gruppen 3 und 4 bei $\Delta L=12$ dB	61
6.1. Verlauf des 2AFC 1-up 1-down Verfahrens	67
6.2. Schematische Darstellung der HKT	77
6.3. THK harmonisch komplexer Töne ohne Energieausgleich	78
6.4. Schematische Darstellung der VHKT	82
6.5. THK für verschobene harmonische komplexe Töne (VHKT)	83
6.6. Einfluss des Abstandes zweier Sinustöne auf die THK	84
6.7. Einfluss des absoluten Darbietungspegel auf die THK	87
6.8. THK von Schmalbandrauschen	89

6.9. THK amplitudenmodulierter Töne	92
6.10. THK frequenzmodulierter Töne	94
6.11. Einfluss der Tonfrequenz auf die THK	95
6.12. Einfluss der Tonfrequenz auf die THK II	97
6.13. Gemessene und nominelle THK über der Frequenz	98
7.1. Schematische Darstellung der harmonischen Tonkomplexe	105
7.2. Mithörschwelle für die HKTs	107
7.3. Tonhaltigkeit der HKTs für Messungen 1, 3, 5 und 7	114
7.4. Tonhaltigkeit der HKTs für Messungen 2, 4, 6 und 8	117
7.5. Gegenüberstellung der Messungen 1 und 2, 3 und 4, 5 und 6, sowie 7 und 8	118
7.6. Übersicht der Ergebnisse Messungen 1–8	120
7.7. Transformierte THK-Daten für Blocked Tracks	124
7.8. Gemessene und modellierte Daten für die HKTs „ohne Energie- ausgleich“	127
7.9. Gemessene und modellierte Daten für die Blocked Tracks	128
7.10. Gemessene und modellierte Daten für die HKTs, Messungen 1–8	129
A.1. Leistungsdichtespektrum für verschiedene Rauschtypen	136
A.2. Verdeckungsmaß	138
C.1. GUI für das Repertory-Grid: Evokationsphase	142
C.2. GUI für das Repertory-Grid: Rangreihe	143
C.3. GUI für das Repertory-Grid: quasikontinuierliche Skala	144
D.1. Zeitlicher Verlauf der THK für die Orientierungsschalle	146
F.1. Mittelwerte der THK für harmonisch komplexe Töne ohne Ener- gieausgleich	158
F.2. Mittelwerte der THK für verschobene harmonisch komplexe Töne	160
F.3. Mittelwerte der THK über den Abstand zweier Sinustöne	161
F.4. Mittelwerte der THK für den absoluten Darbietungspegel	162
F.5. Mittelwerte der THK für die Breite des Schmalbandrauschens	163
F.6. Mittelwerte der THK für amplitudenmodulierte Töne	164
F.7. Mittelwerte der THK für frequenzmodulierte Töne	165
F.8. Mittelwerte der THK über der Tonfrequenz	166
G.1. Gemessene Pegeldifferenzen $L_D(k)$ für Messungen 2, 4, 6 und 8	170
G.2. Korrigierte Pegeldifferenzen $L_{D_{kor}}(k)$ für Messungen 2, 4, 6 und 8	171
G.3. Rohdaten der THK für die Messung mit Blocked Tracks	172
G.4. Transformierte THK-Daten für Blocked Tracks	173

Tabellenverzeichnis

2.1. Tonzuschlag K_T in Abhängigkeit von ΔL	5
3.1. Stimuli für die Untersuchung der maßgeblichen Differenz	20
3.2. Evokationsphase: Konstruktpaare	22
4.1. Übersicht der Probanden	33
5.1. Übersicht der Schalle	51
5.2. Schalle zur Untersuchung der maßgeblichen Differenz	52
5.3. Schalle zur Untersuchung der Frequenz	57
6.1. Tonzuschlag für das Referenzsignal	72
6.2. Paarweise Wilcoxon-Tests für die HKT-Stimuli	79
6.3. Wilcoxon-Tests für die Darbietungspegel	86
7.1. Stimulus-Unterschiede für die Mithörschwelle	108
7.2. Messdesigns für die Messungen der THK	111
7.3. Signifikante Unterschiede des Messdesigns bei Bewertung als Teststimuli	115
7.4. Stimulus-Unterschiede bei Bewertung als Teststimuli	115
7.5. Stimulus-Unterschiede bei Bewertung der Vergleichsstimuli	116
7.6. Stimulus-Unterschiede über alle Messungen	122
7.7. Mittelwerte der HKTs über alle Messungen	122
7.8. Gemessene THK mit Blocked Tracks	123
7.9. Gemessene THK für die HKTs „ohne Energieausgleich“	125
7.10. Mit Blocked Tracks gemessene THK für die HKTs	125
7.11. THK der Messungen 1–8	125
B.1. Umweltgeräusche von Pompetzki	139
D.1. Äquivalenter Dauerschallpegel der Orientierungsschalle	145
D.2. Kategoriale Urteile für die Orientierungsschalle	147
D.3. Häufigkeit der kategorialen Urteile	148
D.4. Kumulierte relative Häufigkeiten der kategorialen Urteile	149
D.5. Abszissenwerte der Standardnormalverteilung	149

E.1. Berechnete THK für die Stimuli	152
F.1. Verwendete Stimuli	155
F.2. Maßgebliche Differenzen für die Teiltöne der HKTs	156
F.3. T-Tests für die HKT-Stimuli ohne Energieausgleich	159
F.4. T-Tests für die Abstände der Sinustöne	160
F.5. T-Tests für die Darbietungspegel	161
F.6. T-Tests für die Breite von Schmalbandrauschen	163
F.7. T-Tests für amplitudenmodulierte Töne	164
F.8. T-Tests für frequenzmodulierte Töne	165
F.9. T-Tests für die Tonfrequenzen (Vergleichsstimulus 700 Hz)	167
F.10. T-Tests für die Tonfrequenzen (Vergleichsstimulus 1850 Hz)	168
G.1. Messwerte für die Messung mit Blocked Tracks	169
G.2. Transformierte Messwerte für die Blocked Tracks	172
G.3. Absolute Messwerte für die Blocked Tracks	173

Literaturverzeichnis

- [BD06] BORTZ, Jürgen ; DÖRING, Nicola: *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. 4., überarbeitete Auflage. Heidelberg : Springer Medizin Verlag, 2006. – ISBN 13 978-3-540-33305-0
- [BEPW00] BACKHAUS, Klaus ; ERICHSON, Bernd ; PLINKE, Wulff ; WEIBER, Rolf: *Multivariate Analysemethoden: eine anwendungsorientierte Einführung*. 9., überarbeitete und erweiterte Auflage. Berlin [u.a.] : Springer, 2000 (Springer-Lehrbuch). – ISBN 3-540-67146-3
- [BFP97] BUUS, Soren ; FLORENTINE, Mary ; POULSEN, Torben: Temporal integration of loudness, loudness discrimination, and the form of the loudness function. In: *J. Acoust. Soc. Am.* 101 (1997), Februar, Nr. 2, S. 669–680
- [BLB00] BORTZ, Jürgen ; LIENERT, Gustav A. ; BOEHNKE, Klaus: *Verteilungsfreie Methoden in der Biostatistik*. 2. Auflage. Berlin, Heidelberg, New York : Springer-Verlag, 2000. – ISBN 3-540-67590-6
- [Bor93] BORTZ, Jürgen: *Statistik für Sozialwissenschaftler*. 4., vollständig überarbeitete Auflage. Berlin, Heidelberg, New York : Springer, 1993. – ISBN 3-540-56200-1
- [BSFZ86] BUUS, Soren ; SCHORER, Edwin ; FLORENTINE, Mary ; ZWICKER, Eberhard: Decision rules in detection of simple and complex tones. In: *J. Acoust. Soc. Am.* 80 (1986), Nr. 6, S. 1646–1657
- [BT59] BLACKMAN, R.B. ; TUKEY, John W.: *The measurement of power spectra. From the point of view of communications engineering*. New York: Dover Publications, 1959
- [Cat66] CATTELL, Raymond B.: The scree test for the number of factors. In: *Multivariate Behavioral Research* 1 (1966), Nr. 2, S. 245–276
- [CH93] COMMANDEUR, J. J. F. ; HEISER, W. J.: Mathematical derivations in the proximity scaling (PROXSCAL) of symmetric data matrices. Leiden, The Netherlands : University of Leiden: Department of Data Theory, 1993 (Research Report RR-93-04). – Forschungsbericht

- [Dan95] DANIEL, Peter: *Berechnung und kategoriale Beurteilung der Rauigkeit und Unangenehmheit von synthetischen und technischen Schallen*, Universität Oldenburg, Diss., 1995
- [DIN89] *DIN 45573 Teil 4: Lautsprecher-Prüfverfahren; Hörtest; Paarvergleiche*. 1989
- [Dud00] DUDENREDAKTION (Hrsg.): *Die deutsche Rechtschreibung*. 22., völlig neu bearbeitete und erweiterte Auflage. Mannheim : Bibliographisches Institut & F. A. Brockhaus, 2000. – ISBN 3-411-04012-2
- [Fas97] FASTL, Hugo: Pitch Strength and Frequency Discrimination for Noise Bands or Complex Tones. In: PALMER, A. R. (Hrsg.) ; REES, A. (Hrsg.) ; SUMMERFIELD, A.Q. (Hrsg.) ; MEDDIS, R. (Hrsg.): *Psychophysical and Physiological Advances in Hearing*. London : Whurr, 1997 (Proceedings of the 11th International Symposium on Hearing), S. 209–215
- [FG50] FLETCHER, Harvey ; GALT, Rogers H.: The Perception of Speech and its Relation to Telephony. In: *J. Acoust. Soc. Am.* 22 (1950), März, Nr. 2, S. 89–151
- [Fra10] FRAUNHOFER IDMT, PROJECT GROUP HEARING, SPEECH AND AUDIO TECHNOLOGY: *Speech Intelligibility Prediction Toolbox*. Version 1.0. März 2010
- [Fru04] FRUHMANN, Markus: On the pitch strength of bandbass noises. In: *ICA 2004*, 2004, S. 1791–1794
- [FZ07] FASTL, Hugo ; ZWICKER, Eberhardt: *Psychoacoustics: Facts and Models*. 3. Edition. Berlin, Heidelberg : Springer, 2007. – 462 S.
- [Gel98] GELFAND, Stanley A.: *Hearing: an introduction to psychological and physiological acoustics*. 3. edited, revised and expanded Edition. New York : Dekker, 1998. – ISBN 0-8247-0143-7
- [GH97] GROSE, John H. ; HALL, Joseph W.: Multiband detection of energy fluctuations. In: *J. Acoust. Soc. Am.* 102 (1997), August, Nr. 2, S. 1088–1096
- [GML59] GREEN, David M. ; MCKEY, Mary J. ; LICKLIDER, J. C. R.: Detection of a pulsed sinusoid as a function of frequency. In: *J. Acoust. Soc. Am.* 31 (1959), Nr. 11, S. 1446–1452
- [Han10] HANSEN, Hans: *Tone-noise dichotomy : investigating tonal content magnitude & pitch strength*, Universität Oldenburg, Diss., 2010

- [Har05] HARTMANN, William M.: *Signals, Sound and Sensation*. Corr. 5. print. Springer Science+Business Media, 2005. – ISBN 1–56396–283–7
- [Hea98] HEAD ACOUSTICS: *Artemis, Multikanal-Analyse-Software*. Version 11.5.1. Head Acoustics, Ebertstraße 30a, 52134 Herzogenrath, 1998
- [Hea10] HEAD ACOUSTICS: *HEAD Application Note: Psychoakustische Analysen in Artemis I - Lautheits- und Schärfe-Berechnung in Artemis*. Head Acoustics, Ebertstraße 30a, 52134 Herzogenrath. http://head-acoustics.de/downloads/de/application_notes/Psychoakustische_Analysen_I_03_10d.pdf. Version: März 2010
- [HEKZ08] HELLBRÜCK, Jürgen ; ELLERMEIER, Wolfgang ; KOHLRAUSCH, Armin ; ZEITLER, Alfred: *Kompendium zur Durchführung von Hörversuchen in Wissenschaft und industrieller Praxis (Entwurf)*. Version: September 2008. http://www.dega-akustik.de/publikationen/Kompendium_Hoerversuche_2008.pdf. Deutsche Gesellschaft für Akustik e.V., Voltastraße 5, 13355 Berlin
- [Hel93] HELLBRÜCK, Jürgen: *Hören - Physiologie, Psychologie und Pathologie*. 1. Auflage. Hogrefe Göttingen, 1993. – ISBN 3–8017–0491–2
- [HN02] HELLWEG, Robert D. ; NOBILE, Matthew A.: Modification to procedures for determining prominent discrete tones. In: *Internoise 2002*, 2002
- [Kel55] KELLY, George A.: *The Psychology of Personal Constructs*. Bd. 1. Norton and Company Inc. New York, 1955
- [Kru05] KRUSE, Peter: Weiche Faktoren werden zu harten Fakten. In: *planung&analyse. Zeitschrift für Marktforschung und Marketing* (2005), Nr. 3, S. 2–6
- [Lev71] LEVITT, H.: Transformed Up-Down Methos in Psychoacoustics. In: *J. Acoust. Soc. Am.* 49 (1971), Nr. 2, S. 467–477
- [Mar56] MARILL, Thomas: Detection theory and psychophysics. Research Laboratory of Electronics, Massachusetts Institute of Technology, 1956 (319). – Forschungsbericht
- [Med10] MEDIAMANUAL.AT: *Denotation und Konnotation*. <http://www.mediamanual.at/mediamanual/workshop/kommunikation/bedeutung/denotation01.php>. Version: 2010. – [Online; Stand 03. Dezember 2010]

- [MRB07] MENZEL, Frank ; ROSENBERGER, Matthias ; BUVE, Janko: Emotionale, intuitive und rationale Konstrukte verstehen. In: *Personalführung* 4 (2007)
- [OSB04] OPPENHEIM, Alan V. ; SCHAFER, Ronald W. ; BUCK, John R.: *Zeitdiskrete Signalverarbeitung*. Pearson Studium, 2004. – ISBN 3-8273-7077-9
- [Pio03] PIORR, Detlef: Erfahrungen mit der DIN 45681(Entwurf November 2002). In: *Fortschritte der Akustik: Plenarvorträge und Fachbeiträge der 28. Deutschen Jahrestagung für Akustik DAGA '02, Bochum, 2003*, S. 596–597
- [PK96] PÜSCHEL, D. ; KOCH, R.: *SI – Signalverarbeitung interaktiv*. Version 11.5.1. DPI-Press, Juni 1996
- [Pom98] POMPETZKI, W.: Vergleich der Tonhaltigkeit nach DIN 45 681 mit subjektiven Bewertungen. In: SILL, Albert (Hrsg.): *Plenarvorträge und Fachbeiträge der 24. Deutschen Jahrestagung für Akustik, DAGA '98*. Universität Oldenburg, Physik/Akustik, D-26111 Oldenburg : Deutsche Gesellschaft für Akustik e.V., 1998. – ISBN 3-9804568-3-8, S. 224–225
- [PRA01] PÜSCHEL, D. ; RÖSLER, S. ; ALPHEI, H.: Zur Tonhaltigkeit von Geräuschen. In: *Fortschritte der Akustik : Plenarvorträge und Fachbeiträge der 27. Deutschen Jahrestagung für Akustik DAGA 2001, Hamburg, 2001*
- [Rie00] RIEDEL, Volker: *Tonhaltigkeit (Kapitel 7) im "Skript zum Blockpraktikum Psychoakustik und auditorische Signalverarbeitung". Praktikum durchgeführt 21.09.–02.10.1998 an der Universität Oldenburg, Betreuer Matthias Vormann und Jesko Verhey, verantwortlich Birger Kollmeier und Volker Hohmann, AG Medizinische Physik am FB 8, Physik*. 2000
- [Rit67] RITSMA, Roelof J.: Frequencies dominant in the Perception of the Pitch of Complex Sounds. In: *J. Acoust. Soc. Am.* 42 (1967), Nr. 1, S. 191–198
- [RR99] ROBERTS, Maxwell J. ; RUSSO, Riccardo: *A student's guide to analysis of variance*. Routledge London [u.a.], 1999. – ISBN 0-415-16564-4 – 0-415-16565-2

- [Sag01] SAGEMÜHL, Dirk: Überarbeitung der E DIN 45681 (Tonhaltigkeit). In: *Fortschritte der Akustik : Plenarvorträge und Fachbeiträge der 27. Deutschen Jahrestagung für Akustik DAGA 2001, Hamburg, 2001*
- [SC93] SCHEER, Jörn W. ; CATINA, Ana: *Einführung in die Repertory Grid-Technik*. Bd. 1: *Grundlagen und Methoden*. Bern, Göttingen, Toronto, Seattle : Verlag Hans Huber, 1993
- [Sch61] SCHARF, Bertram: Loudness Summation under Masking. In: *J. Acoust. Soc. Am.* 33 (1961), April, Nr. 4, S. 503–511
- [Sch92] SCHEER, Jörn W.: Psychologie der persönlichen Konstrukte und Repertory Grid-Technik. In: SCHMIDT, L. R. (Hrsg.): *Psychologische Aspekte medizinischer Maßnahmen. Jahrbuch der medizinischen Psychologie* Bd. 7. Berlin, Heidelberg : Springer-Verlag, 1992, S. 273–290
- [SG49] SCHAFER, T.H. ; GALES, R. S.: Auditory masking of multiple tones by random noise. In: *J. Acoust. Soc. Am.* 21 (1949), S. 392–398
- [Six67] SIXTL, Friedrich: *Meßmethoden der Psychologie*. Weinheim a.d. Bergstraße : Verlag Julius Beltz, 1967
- [SPS02] SPSS INC.: *SPSS for Windows*. Version 11.5.1. SPSS Inc. Headquarters, 233 S. Wacker Drive, 11th floor, Chicago, Illinois 60606, November 2002
- [SPS09] SPSS INC.: *PROXSCAL*. Version: 2009. <http://support.spss.com/ProductsExt/SPSS/Documentation/Statistics/algorithms/14.0/proxscal.pdf>. SPSS Inc. Headquarters, 233 S. Wacker Drive, 11th floor, Chicago, Illinois 60606. Online im Internet. – (Stand 20.11.2009)
- [SS03] SAGEMÜHL, Dirk ; SCHMIDT, Lothar: Tonhaltigkeit - Stand der E DIN 45681 einschließlich weiterer Ergänzungen. In: *Fortschritte der Akustik: Plenarvorträge und Fachbeiträge der 29. Jahrestagung für Akustik DAGA 03, Aachen*. Oldenburg : DEGA, 2003, S. 594–595
- [SW94] SPRINGER, Nils ; WEBER, Reinhard: Existenzbereich der R-Rauhigkeit. In: *Fortschritte der Akustik, DAGA 1994* DPG-GmbH, Bad Honnef, 1994, S. 1221–1224
- [Ter98] TERHARDT, Ernst: *Akustische Kommunikation: Grundlagen mit Hörbeispielen*. 1. Auflage. Berlin, Heidelberg, New York : Springer, 1998. – ISBN 3–540–63408–8

- [The06] THE MATHWORKS, INC.: *Matlab R2006b*. August 2006. – Version 7.3.0.267
- [THK92] Norm DIN 45681:1992-01E Januar 1992. *Bestimmung der Tonhaltigkeit von Geräuschen und Ermittlung eines Tonzuschlages für die Beurteilung von Geräuschmissionen - Entwurf*. – Vertrieb durch Beuth Verlag GmbH, Berlin
- [THK05] Norm DIN 45681:2005-03 März 2005. *Akustik - Bestimmung der Tonhaltigkeit von Geräuschen und Ermittlung eines Tonzuschlages für die Beurteilung von Geräuschmissionen*. – Vertrieb durch Beuth Verlag GmbH, Berlin
- [THK06] Norm DIN 45681:2006-08 August 2006. *Akustik - Bestimmung der Tonhaltigkeit von Geräuschen und Ermittlung eines Tonzuschlages für die Beurteilung von Geräuschmissionen, Berichtigungen zu DIN 45681:2005-03, mit CD*. – Vertrieb durch Beuth Verlag GmbH, Berlin
- [TSS82] TERHARDT, Ernst ; STOLL, Gerhard ; SEEWANN, Manfred: Algorithm for Extraction of Pitch and Pitch Saliency from Complex Tonal Signals. In: *J. Acoust. Soc. Am.* 71 (1982), März, Nr. 3, S. 679–688
- [Tuk77] TUKEY, John W.: *Exploratory data analysis*. Addison-Wesley Series in Behavioral Science: Quantitative Methods, Reading, Mass.: Addison-Wesley, 1977
- [Ver99] VERHEY, Jesko L.: *Psychoacoustics of spectro-temporal effects in masking and loudness perception*. Bibliotheks- und Informationssystem der Universität Oldenburg, Universität Oldenburg, Diss., 1999
- [Vor95] VORMANN, Matthias: *Psychoakustische Modellierung der virtuellen Tonhöhe*. Fachbereich Physik, Carl von Ossietzky-Universität Oldenburg, Diplomarbeit, September 1995
- [VU08] VERHEY, Jesko L. ; UHLEMANN, Michael: Spectral loudness summation for sequences of short noise bursts. In: *J. Acoust. Soc. Am.* 123 (2008), S. 925–934
- [VVM⁺99] VORMANN, Matthias ; VERHEY, Jesko L. ; MEIS, Markus ; MEL-LERT, Volker ; SCHICK, August: Factors Influencing the Subjective Rating of Tonal Components in Noise. In: *Collected Papers from the Joint Meeting 'Berlin 99': 137th regular meeting of the Acoustical Society of America, 2nd convention of the EAA: Forum Acusticum*

- *integrating the 25th German Acoustics DAGA Conference*. Universität Oldenburg, Physik/Akustik, D-26111 Oldenburg : Deutsche Gesellschaft für Akustik e.V., 1999

- [VVMS00] VORMANN, Matthias ; VERHEY, Jesko L. ; MELLERT, Volker ; SCHICK, August: Subjective Rating of Tonal Components with an Adaptive Procedure. In: SCHICK, August (Hrsg.) ; MEIS, Markus (Hrsg.) ; RECKHARDT, Carsten (Hrsg.): *Contributions to Psychological Acoustics: Results of the Eighth Symposium on Psychological Acoustics*. 1. Edition. Postfach 2541, 26015 Oldenburg : Bibliotheks- und Informationssystem der Universität Oldenburg, 2000 (Akustik und Schwingungstechnik), S. 145–153
- [Wik09] WIKIPEDIA: *Wahrnehmung* — *Wikipedia, Die freie Enzyklopädie*. <http://de.wikipedia.org/w/index.php?title=Wahrnehmung&oldid=63742863>. Version: 2009. – [Online; Stand 23. September 2009]
- [Wik10] WIKIPEDIA: *Grounded Theory* — *Wikipedia, Die freie Enzyklopädie*. http://de.wikipedia.org/w/index.php?title=Grounded_Theory&oldid=79882144. Version: 2010. – [Online; Stand 3. Dezember 2010]
- [ZF90] ZWICKER, Eberhardt ; FASTL, Hugo: *Psychoacoustics*. 1. Edition. Berlin, Heidelberg : Springer Verlag, 1990 (Springer Series in Information Sciences 22)
- [ZG04] ZIMBARDO, Philip G. ; GERRIG, Richard J.: *Psychologie*. 16. aktualisierte Auflage. Pearson Studium, München, 2004. – ISBN 3–8273–7056–6
- [ZT80] ZWICKER, Eberhardt ; TERHARDT, Ernst: Analytical expressions for critical band rate and critical bandwidth as a function of frequency. In: *J. Acoust. Soc. Am.* 68 (1980), November, Nr. 5, S. 1523–1525
- [Zwi61] ZWICKER, Eberhardt: Subdivision of the Audible Frequency Range into Critical Bands (Frequenzgruppen). In: *J. Acoust. Soc. Am.* 33 (1961), Februar, Nr. 2, S. 248
- [Zwi82] ZWICKER, Eberhardt: *Psychoakustik*. 1. Auflage. Berlin; Heidelberg; New York : Springer, 1982 (Hochschultext). – ISBN 3–540–11401–7

Danksagung

Mein herzlicher Dank gilt allen, die direkt und indirekt am Zustandekommen dieser Arbeit beteiligt waren.

Insbesondere danke ich Prof. Dr. Jesko Verhey und Prof. Dr. Volker Mellert für die wohlwollende Betreuung und Begutachtung der vorliegenden Arbeit sowie für ihre stetige Hilfe und Geduld.

Beim „Tontreffen“ hatte ich mit Dr. Reinhard Weber, Dr. Hans Hansen, Stephan Töpken und Arne Oetjen kritische Zuhörer, die durch Ihre Anregungen und Nachfragen meinen Ideen und Forschungen ständigen Antrieb gegeben haben.

Prof. August Schick, Dr. Maria Klatte und Dr. Markus Meis lieferten in stets konstruktiven Diskussionen wertvolle Hinweise für die Entwicklung und Darstellung der subjektiven Messungen der Tonhaltigkeit, insbesondere des methodischen Teils der perceptuellen Aspekte der Tonhaltigkeit.

Wesentliche Teile dieser Arbeit wurden angefertigt im Rahmen des von der Deutschen Forschungsgemeinschaft geförderten interdisziplinären Graduiertenkollegs „Psychoakustik: Auditorische Verarbeitung, Schallbewertung und Schallwirkung - Modellierung grundlegender Gehöreigenschaften für technische Anwendungen und klinischen Einsatz“. Ich danke der DFG für die Förderung und allen Kollegen und Professoren des Graduiertenkollegs für die Möglichkeit zur interdisziplinären Forschung und internationalem Wissensaustausch.

Für die Durchführung der vorgestellten Versuche konnte ich die Labore und Geräte der Arbeitsgruppe Akustik des Fachbereiches Physik an der Carl von Ossietzky Universität Oldenburg nutzen.

Darüber hinaus danke ich Dr. Andreas Liebl, Dr. Carsten Sommer und Dr. Peter Daniel für fruchtbare Diskussionen und regen Ideenaustausch, nicht nur die Tonhaltigkeit betreffend.

Weiterhin bedanke ich mich bei Stefanie Pfefferle und Dania Molitor, die die Lesbarkeit der vorliegenden Arbeit verbessert und für die Einhaltung eines Mindeststandards an orthographischen Regeln gesorgt haben.

Abschließend möchte ich meinen Dank allen nicht explizit genannten Kollegen und Freunden vom Hörzentrum Oldenburg, der Hörtech, der Medi, dem Itap, der Arbeits-, Umwelt- und Gesundheitspsychologie an der KU Eichstätt, dem Akustikbüro Oldenburg, meiner Familie und insbesondere allen Probanden aussprechen: „Vielen Dank, ohne Euch wäre das nicht möglich gewesen!“

Kurzlebenslauf

Name: Matthias Vormann
Geburtsdatum: 20. Juli 1968
Geburtsort: Hagen/Westfalen
Staatsangehörigkeit: deutsch

Schulbildung

1975 — 1979 Grundschule Breckerfeld
1979 — 1985 St. Jacobus Realschule Breckerfeld
1985 — 1988 Abitur: Anne Frank Gymnasium Halver

Studium

01.10.89 — 30.09.92 Physikstudium an der Universität Dortmund
01.10.92 — 09.11.95 Physikstudium an der CvO Universität Oldenburg
Abschluss: Diplom-Physiker

Berufstätigkeit

01.01.96 — 30.04.98 Verschiedene Tätigkeiten an der Carl von Ossietzky Universität Oldenburg
01.05.98 — 31.03.01 Promotionsstipendium im Graduiertenkolleg „Psychoakustik“ an der CvO Universität Oldenburg
01.04.01 — 31.03.03 Diverse Tätigkeiten für verschiedene Arbeitgeber
01.04.03 — 31.07.05 Wissenschaftlicher Mitarbeiter an der katholischen Universität Eichstätt-Ingolstadt
01.08.05 — 15.12.08 Wissenschaftlicher Mitarbeiter am Hörzentrum Oldenburg GmbH, Oldenburg im Bereich Audiologie und Projekte
01.04.09 — 15.07.10
ab 01.01.2011
16.12.08 — 31.03.09 Projektmitarbeit als wissenschaftlicher Mitarbeiter am
16.07.10 — 31.12.10 Institut für Physik an der CvO Universität Oldenburg

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst und nur die angegebenen Quellen und Hilfsmittel verwendet habe.

Oldenburg, den 10. Dezember 2010

_____ (Matthias Vormann)