Technische Universität Berlin
Fakultät I - Geisteswissenschaften
Institut für Sprache und Kommunikation
Fachgebiet Musikwissenschaft

Magisterarbeit

# DIMENSIONEN DER KLANGFARBE

## Skalierung von Instrumentalklängen unterschiedlicher Tonhöhe

Sven-Amin Lembke
[redacted]

Erstgutachterin: Prof. Dr. Helga de la Motte-Haber
Zweitgutachterin: Prof. Dr. Brigitte Schulte-Fortkamp

# ERKLÄRUNG

Die selbständige Anfertigung versichere ich an Eides statt. Alle Tabellen und Abbildungen wurden selber erstellt und beruhen auf eigenen empirischen Erhebungen.

Berlin, den 8. September 2006

[redacted]

Sven-Amin Lembke

# INHALTSVERZEICHNIS


## 1 - KLANGFARBE: EINE EINFÜHRUNG

1.1 - Fragestellung: Die Dimensionalität der Klangfarbe 1
1.2 - Multidimensionale Skalierung der Klangfarbe 2
1.2.1 - Zusammenfassung bisheriger Forschungsergebnisse 3
1.2.2 - Studie von Handel & Erickson (2004) 5
1.3 - Zielsetzung dieser Arbeit 7

## 2 - AUSWAHL DER INSTRUMENTE & TONHÖHEN

2.1 - Auswahl der Instrumente 10
2.2 - Auswahl der Tonhöhen 11
2.3 - Die Stimuli und ihre Standardisierung 12
2.3.1 - Aufnahme der Stimuli 12
2.3.2 - Standardisierung der Stimuli 15
2.4 - Instrumentenspezifische Charakteristika über die gewählten Tonhöhen 18
2.4.1 - Klarinette 18
2.4.2 - Horn 22
2.4.3 - Baßflöte 28
2.5 - Vergleich der Instrumentalklänge 31
2.5.1 - Unterschiede im Spektrum 32
2.5.2 - Unterschiede im Einschwingverhalten 33

## 3 - VERSUCHSPLAN & VERSUCHSDURCHFÜHRUNG

3.1 - Versuchsplan 35
3.1.1 - Zusammensetzung und Darbietungsform der Stimuli-Paare 35
3.1.2 - Bewertungsskalierung 36
3.2 - Versuchsdurchführung 37
3.2.1 - Versuchspersonen 38
3.2.2 - Technische Umsetzung des Versuchs 39

## 4 - AUSWERTUNG

4.1 - Multidimensionale Skalierung und ihre Terminologie 40
4.1.1 - Eingabe- und Ausgabegrößen einer MDS 40
4.1.2 - Der Zusammenhang zwischen Proximitäten $p_{ij}$ und Distanzen $d_{ij}$ 41
4.1.3 - Handhabung gleicher Proximitäten $p_{ij}$ in einer ordinalen MDS 42
4.1.4 - Bewertung der Qualität eines MDS-Ergebnisses 43
4.2 - Aufbereitung und Überprüfung der erhobenen Proximitäten 45
4.2.1 - Überprüfung der Matrix-Symmetrie-Voraussetzung 45
4.2.2 - Die erhobenen Proximitäten $p_{ij}$ 46

4.3 - Durchführung der MDS 47
4.3.1 - Die Algorithmen ALSCAL und PROXSCAL im Vergleich 47
4.3.2 - Die Parameterwahl in PROXSCAL 49
4.4 - Präsentation zwei- und dreidimensionaler PROXSCAL-Ergebnisse 51
4.4.1 - Zweidimensionale MDS 51
4.4.2 - Dreidimensionale MDS 53
4.4.3 - Zwei- und dreidimensionale Lösungen im Vergleich 55

5 - DEUTUNG DER DIMENSIONEN

5.1 - Deskriptive Betrachtung der Konstellation X 59
5.2 - Ansätze zur Deutung der Dimensionen 61
5.3 - Einführende Zusammenfassung der Vergleichsgrößen 62
5.3.1 - Untergruppe 3 63
5.3.2 - Untergruppe 1 64
5.3.3 - Untergruppe 2 67
5.3.4 - Korrelationen zwischen den Vergleichsgrößen und Dimensionen 69
5.4 - Quantitative und qualitative Deutung der dimensionalen Struktur 70
5.4.1 - Vorbetrachtung ‚unerwünschter' Einflußgrößen 70
5.4.2 - Diskussion und Deutung von Dimension 1 73
5.4.3 - Diskussion und Deutung von Dimension 2 77
5.4.4 - Diskussion und Deutung von Dimension 3 80
5.4.5 - Zusammenfassung 84

6 - DISKUSSION & VERGLEICH DES MDS-ERGEBNISSES

6.1 - Vergleich des vorliegenden Ergebnisses mit Handel & Erickson (2004) 86
6.2 - Diskussion der Unterschiede zwischen den Versuchen 87
6.2.1 - Unterschiede hinsichtlich der Instrumente 88
6.2.2 - Unterschiede hinsichtlich der Tonhöhe 89
6.2.3 - Weitere Unterschiede 93
6.3 - Zusammenfassung und Ausblick 95

LITERATURVERZEICHNIS 99

ANHANG A 101

# 1 - KLANGFARBE: EINE EINFÜHRUNG

## 1.1 - Fragestellung: Die Dimensionalität der Klangfarbe

Die uns bekannte abendländische Musik kann seit der Entwicklung des heutzutage gebräuchlichen Notationssystems in vielerlei Hinsicht problemlos zu Papier gebracht werden. Die wichtigsten Eigenschaften eines musikalischen Tons bzw. Klangs lassen sich dadurch festlegen: Tonhöhen lassen sich in vertikaler, Tondauern in horizontaler Ausrichtung und die Lautstärke entsprechend einer Ordinalfolge von Symbolen notieren. Für alle dieser Eigenschaften läßt sich dies aus demselben Grund problemlos bewerkstelligen: ihrer im Grunde eindimensionalen Natur. Alle von ihnen erstrecken sich auch auf der semantischen Ebene der Sprache über jeweils eindimensionale Kontinuen: die Tonhöhe zwischen tief und hoch, die Tondauer zwischen kurz und lang, sowie die Lautstärke zwischen leise und laut. Diese Komponenten lassen sich aus dem gleichen Grund auch als physikalische Größen formulieren bzw. skalieren; selbst wenn es in der konkreten Ausführung abweichende Ansätze gibt.

Im Unterschied dazu läßt sich für die Klangfarbe – als beiläufige, gleichzeitig jedoch auch untrennbare Komponente der Tonwahrnehmung – kein eindimensionales Kontinuum finden, welches sie angemessen beschreiben könnte. Weder in der semantischen Ebene von Sprache: die Klangfarbe ist nicht nur zwischen hell und dunkel oder scharf und stumpf oder lieblich und brachial zu beschreiben. Noch in den physikalisch-akustischen Eigenschaften von Klängen, in denen sie bezüglich des Frequenzgehalts sowie der zeitlichen Gestalt eines Klangs charakterisiert wird. Selbst aus dem Notentext wird nicht klar, um welche Klangfarbe es sich konkret handelt. Vielmehr ergibt sie sich aus einem Zusammenwirken des vorgeschriebenen Instruments, sowie den anderen Komponenten Tonhöhe, Lautstärke und Dauer. Somit ist die Klangfarbe schwer faßbar, vielschichtig beeinflußt und sie wirft sehr viele Unklarheiten auf. Einigkeit herrscht jedoch in einer wesentlichen Charakteristik: der multidimensionalen Natur der Klangfarbe.[1]

Dieser komplexe Sachverhalt verleitet geradezu eine allumfassende, möglichst konzise Definition für die Klangfarbe zu suchen. Eine viel zitierte Definition der Klangfarbe (*timbre*) wurde durch das American National Standards Institute (ANSI) formuliert: „Timbre is that attribute of auditory sensation in terms of which a subject can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar.“[2] Bezüglich dieser Definition gibt es einige Übereinstimmung darin, daß sie eigentlich nur eine Definition dafür ist, was die Klangfarbe gerade nicht ist.[3] Doch obwohl Tonhöhen- und Lautstärkeunterschiede eigenständige Komponenten der Tonwahrnehmung sind, kann nicht ganz abgestritten werden, daß sie auch einen Anteil zur Klangfarbe beitragen. Außerdem weist Bregman darauf hin, daß für geräuschhafte bzw. wenig tonale Klänge ohne klar identifizierbare Tonhöhe, die Bestimmung der Klangfarbe nach der ANSI-Definition praktisch unmöglich gemacht wird.[4] Auch gibt es Ansätze, welche bereits anerkennen, daß sich mit der Tonhöhe und Lautstärke auch die Klangfarbe eines Instruments ändert. Handel & Erickson schlagen diesen Weg

1 vgl. Licklider (1951), Plomp (1976), Dowling & Harwood (1986), Butler (1992), Hajda et al. (1997).
2 ANSI (1973). S. 56.
3 vgl. Plomp (1976), Dowling & Harwood (1986), Bregman (1990), Hajda et al. (1997), Reuter (2005).
4 Bregman (1990). S. 92.

ein und postulieren für ein Instrument einen ganzen Satz an Klangfarben, welcher sich nach einzelnen Tonhöhen, sowie Lautstärken aufgliedert.[5] Ganz im Gegenteil besitzt der Begriff *Klangfarbe* im üblichen Sprachgebrauch bzw. -verständnis bezogen auf ein Instrument auch eine verallgemeinernde semantische Bedeutung. So kann im Allgemeinen von der ‚Klangfarbe der Klarinette' die Rede sein, ungeachtet der über die Tonhöhe und Lautstärke stark variierenden Unterschiede in der Klangfarbe.

Weiterhin läßt sich fragen, inwiefern sich Klangfarbe überhaupt absolut ausdrücken läßt. Lassen sowohl verbale Beschreibungen als auch absolute physikalisch-akustische Eigenschaften bereits einen Rückschluß auf die vorliegende Klangfarbe zu? Oder werden wie in einigen Wahrnehmungstheorien relative Bezüge zu anderen Klängen zum Erhalt der Klangfarbe genutzt? Hier würde eine Abhängigkeit gegenüber dem Kontext mit einhergehen, welche die Wandlungsfähigkeit in der Bedeutung der Klangfarbe erklären könnte. Die mehrdimensionale Prägung der Klangfarbe in physikalisch-akustischer oder sprachlich-semantischer Sicht würde diesbezüglich gerade im Verhältnis zu anderen Klängen charakterisiert.

## 1.2 - Multidimensionale Skalierung der Klangfarbe

Für die Erforschung mehrdimensionaler Sachverhalte existiert ein deskriptives, statistisches Verfahren, welches es vermag, eindimensional skalierte Daten über Beziehungen zwischen Objekten auf mehrdimensionale Zusammenhänge aller Objekte zueinander zu untersuchen und in einer geometrischen Darstellung aller Objekte auszudrücken: die sogenannte Multidimensionale Skalierung (MDS). Oftmals basieren die eindimensionalen Eingabedaten selbst bereits auf einer Integration verschiedener Einflüsse, welche in ihrer Eindimensionalität jedoch verborgen bleiben. Eine MDS kann infolge dafür angewandt werden, diese latent wirksamen Einflüsse wieder zum Vorschein zu bringen. Das wohl bekannteste und anschaulichste Beispiel für die Anwendung einer MDS bezieht sich auf die aus Atlanten oder Fernverkehrsbroschüren bekannten Tabellen, in denen Entfernungen zwischen Städten in allen möglichen Kombinationen angegeben sind. Diese lediglich eindimensional skalierte Eingabegröße der Distanz kann mittels einer MDS in einer zweidimensionalen geometrischen Darstellung ausgedrückt werden, welche bestenfalls einer Rekonstruktion der tatsächlichen, geographischen Lage der Städte entspricht.

Während im vorangegangenen Beispiel eine metrische Eingabegröße zu einer ebenfalls metrischen, auf euklidischen Distanzen beruhenden multidimensionalen Abbildung führt, lassen sich annähernd gleichwertige Ergebnisse auch mit ordinalskalierten Eingabedaten realisieren. Dies ermöglicht die MDS auch für die Untersuchung von subjektiven Urteilen einzusetzen, wie beispielsweise in der Soziologie, der Marktforschung oder der Wahrnehmungspsychologie. Hierbei wird die genannte Eingabedimension der Distanz in eine empfundene ‚Nähe' zwischen Objekten umgedeutet, was sich auch als Ähnlichkeit bzw. Unterschiedlichkeit ausdrücken läßt. Auf die notwendigen Urteile der Ähnlichkeit zwischen Objekten kann mittels Korrelationsgrade verschiedener Indikatoren,

5 Handel & Erickson (2004). S. 588.

Verwechslungsmatrizen, sowie direkten Bewertungen von Objektpaaren mit Ratingskalen geschlossen werden. Am Ende läßt sich aus diesen (abgeleiteten oder direkten) subjektiven Urteilen über die Ähnlichkeit zwischen Objekten eine sogar auf Intervallskalenniveau basierte geometrische Repräsentation dieser Objekte generieren, welche in vielen Fällen die ihnen latent zugrundeliegenden Einflüsse bzw. Organisationsstrukturen widerspiegelt.

Als naheliegende Konsequenz wird die MDS in der Musik- bzw. Tonpsychologie vor allem zur Erforschung der Mehrdimensionalität der Klangfarbe herangezogen. An die ersten Studien dieser Art in den 1970er Jahren schlossen sich in der Folgezeit eine Vielzahl weiterer MDS-Untersuchungen zur Klangfarbe an. Es wird daher im anschließenden Teil nicht beabsichtigt, eine vollständige Darstellung dieses Forschungsbereichs wiederzugeben, sondern vielmehr einen für diese Arbeit relevanten Einblick in die Gemeinsamkeiten und Unterschiede von Forschungsergebnissen zu erhalten.

### 1.2.1 - Zusammenfassung bisheriger Forschungsergebnisse

Für diese Arbeit lassen sich die MDS-Studien zur Klangfarbe sinnvollerweise in zwei Gruppen trennen. Während sich eine überwiegende Mehrheit der Studien im Grunde an der ANSI-Definition der Klangfarbe orientiert (und somit losgelöst von Unterschieden der Tonhöhe und Lautstärke agiert), gibt es auch einige Studien, welche den Einfluß der Tonhöhe gezielt als unabhängige Variable in ihr Versuchskonzept aufnehmen. Somit unterscheiden sich die im folgenden vorgestellten Gruppen in der Berücksichtigung bzw. des Verzichts auf eine Standardisierung der Stimuli entsprechend einer konstanten Tonhöhe.

*a) Versuche mit einer Tonhöhe*

Seit ihren Anfängen beschränken sich MDS-Studien zur Klangfarbe überwiegend auf den Vergleich von Stimuli mit gleicher Tonhöhe: Grey (1976), Grey & Moorer (1977), Iverson & Krumhansl (1993), Iverson (1995), McAdams et al. (1995), Kendall et al. (1999), Lakatos (2000). Sie haben alle gemeinsam, daß sich trotz unterschiedlicher Versuchsansätze bzw. -pläne, in ihren dimensionalen Ergebnissen sowohl spektrale als auch temporale Einflußgrößen nachweisen lassen. Während spektrale Eigenschaften in allen Studien eine Rolle spielen, sind temporale Faktoren sehr an die Auswahl und Bandbreite der Hörstimuli gebunden (wie beispielsweise das Vorhandensein sowohl perkussiver als auch kontinuierlich angeregter Klänge). Sie lassen sich daher nur in einigen Fällen zur Erklärung der dimensionalen Struktur nutzen.[6] Bezüglich der Spektraleigenschaften konnte im Zuge der Forschung eine Einzahlwert-Größe formuliert werden, welche es vermag, in vielen dieser Studien mit einer Dimension stark zu korrelieren: der sogenannte *spectral centroid*.[7] Es wird seither versucht, den Einfluß dieser Größe in nahezu jeder MDS-Untersuchung zur Klangfarbe nachzuweisen.

Abgesehen von den oben genannten Gemeinsamkeiten, welche einen allgemeinen Vergleich zwischen den Studien zulassen, kann man in den wenigsten Fällen die MDS-Resultate verschiedener Studien direkt miteinander vergleichen, da es teilweise gravierende Unterschiede in ihrer Versuchsdurchführung

6 Grey (1976), McAdams et al. (1995), Lakatos (2000).
7 Iverson & Krumhansl (1993), McAdams el al. (1995), Kendall et al. (1999), Lakatos (2000).

gibt. Neben den bereits genannten unterschiedlichen temporalen Eigenschaften einzelner Stimuli, sind es die Klangquellen der Stimuli selbst, welche einen direkten Vergleich oftmals nicht zulassen. Die Stimuli der genannten Studien basieren auf einer immensen Bandbreite von Klangerzeungsverfahren: von rein-synthetischen Tonkomplexen, über naturalistisch-synthetische Nachbildungen von Instrumentalklängen, bis hin zu real- akustischen, aufgezeichneten Instrumentalklängen. Hinzu kommt, daß mit der historischen Entwicklung der Tontechnik sich in den Anfängen der digitalen Audiotechnik produzierte Stimuli kaum noch mit der heutzutage gewohnten Klangqualität vergleichen lassen.[8]

Auch in der Standardisierung der Stimuli weichen die Ansätze teilweise erheblich voneinander ab. Während sich Stimuli relativ leicht nach einer festgelegten Tonhöhe standardisieren lassen, gibt es bezüglich der Normierung der Lautstärke grundlegende Unterschiede. So wird in manchen Versuchen lediglich anhand des physikalischen, logarithmierten Pegels normiert.[9] Andere Ansätze berücksichtigen die frequenzabhängige Empfindlichkeit der menschlichen Hörwahrnehmung für Schalldruckpegel und die dadurch bedingte sehr komplexe Bildung einer Lautstärkeempfindung. Sie verzichten auf die Verwendung des wenig aussagekräftigen Pegelwerts zugunsten einer gehörgerechteren Standardisierung. Als Alternative kann beispielsweise eine subjektive Normalisierung der Stimuli durch mehrere Personen vorgenommen werden.[10]

Bezüglich der musikalischen Vorbildung der Versuchsteilnehmer gibt es ebenfalls Unterschiede, wobei manche der Studien gerade diesen Faktor mit zum Untersuchungsgegenstand machen. Während Unterschiede in der musikalischen Vorbildung in den Studien überwiegend nicht zu Unterschieden in den Bewertungen führten[11], gibt es auch Ergebnisse mit einer gegensätzlichen Aussage[12]. Es sei jedoch vorweggenommen, daß eine musikalische Vorbildung aus anderen, später noch genannten Gründen, bei MDS-Untersuchungen mit mehreren Tonhöhen von Vorteil sein könnte.

*b) Versuche mit unterschiedlichen Tonhöhen*

Versuche mit einer Tonhöhe beschränken sich auf Vergleiche zwischen einzelnen Stimuli, ohne jedoch Rückschlüsse auf die tatsächlichen Beziehungen zwischen Instrumenten zuzulassen. Wenn es darum geht die Klangfarbe bzw. Klangfarben eines Instruments zu charakterisieren oder die Trennung zwischen Instrumenten mittels MDS ergründen zu können, reicht die Beschränkung auf eine Tonhöhe nicht aus. Über den Tonraum kann es auf intra-instrumentaler Ebene bereits zu erheblichen klanglichen Unterschieden kommen, welche sonst ignoriert würden. Mit der Hinzunahme mehrerer Tonhöhen wird jedoch auch der Weg frei gemacht, daß ein eventueller Einfluß der Tonhöhe in das Ergebnis Einzug halten kann. Einige Studien, die die Tonhöhe als unabhängige Variable zur

8 Bei Miller & Carterette (1975) wurde für die Stimuli-Wiedergabe eine D/A-Wandlung mit einer Abtastrate von 12,5 kHz verwendet. Grey (1976) und Grey & Moorer (1977) wandelten dagegen schon mit 22,05 bzw. 25 kHz (und 12-Bit Auflösung). Seit nun mehr als 20 Jahren ist inzwischen die *CD-Qualität* mit 44,1 kHz und 16-Bit Minimalstandard. Im Vergleich besitzen die älteren Klänge eine klar minderwertige Klangqualität, was sich sowohl auf die gehörgerechte Abhörsituation als auch die Präzision der physikalisch-akustischen Analysen auswirkt.

9 Iverson (1995), Erickson (2001).

10 Lakatos (2000), Marozeau et al. (2003).

11 Miller & Carterette (1975), Kendall et al. (1999), Lakatos (2000).

12 McAdams et al. (1995).

Untersuchung ihres eventuellen Einflußes beinhalten, seien hier genannt: Miller & Carterette (1975), Erickson (2003), Marozeau et al. (2003), Handel & Erickson (2004).

Ein Einfluß der Tonhöhe läßt sich in allen dieser Studien nachweisen, jedoch mit teilweise grundsätzlich unterschiedlicher Bedeutung für das dimensionale Ergebnis. Während in den meisten Fällen die Tonhöhe eine strukturgebende Stellung einnimmt[13], spielt die Tonhöhe in einem Fall nur eine untergeordnete Rolle[14]. Auch für diese Studien gelten die meisten der unter a) bereits genannten Gemeinsamkeiten in den Ergebnissen, wie auch die Problematik von Unterschieden in der Versuchsdurchführung. Die spektrale Vergleichsgröße *spectral centroid* führt ebenfalls in allen Fällen zu signifikanten Korrelationen mit einer Dimension.

### 1.2.2 - Studie von Handel & Erickson (2004)

In dieser Arbeit dient die Studie von Handel & Erickson (2004) als Grundlage für einen Versuch mit anschließender MDS-Untersuchung. Handel & Erickson nehmen als eine der ersten eine MDS-Untersuchung von Instrumentalklängen unterschiedlicher Tonhöhe vor. Ihre Studie verfolgt in direkter Anlehnung an Pendants mit nur einer Tonhöhe einen eher explorativen Ansatz, ohne wie bei Miller & Carterette oder Marozeau et al. die Tonhöhe als den dominierenden Gegenstand des Versuchs zu wählen. Im diesem Abschnitt soll etwas ausführlicher auf die Durchführung und Ergebnisse dieser Studie eingegangen werden.

Handel & Erickson (2004) gehen von der Klangfarbe als einer „*object representation*“ aus.[15] Danach läßt sich die Klangfarbe nicht allein aus den physikalisch-akustischen Eigenschaften eines einzelnen Klangs beschreiben, sondern sie ist auch mit ihrer Klangquelle selbst, als physisches Objekt, eng verknüpft. Diese ‚Quell-Klangfarbe‘ (*source timbre*) läßt sich jedoch nur bei Vorhandensein von mehreren Klängen unterschiedlicher Tonhöhe bzw. Lautstärke einer Klangquelle extrahieren. Nach Handel & Erickson ist das auch ein Grund, warum in MDS-Studien mit nur einer Tonhöhe den Forschern bislang versagt blieb, im Ergebnis eine wirkliche kategoriale Trennung von Instrumenten bzw. Klangquellen zu erhalten. Denn die Ergebnisse dieser Studien basieren auf Bewertungen, welche ausschließlich (auf einer niedrigen sensorischen Ebene) physikalisch-akustische Unterschiede widerspiegeln.

Daher schlagen Handel & Erickson für die Untersuchung der ‚Quell-Klangfarbe‘ die Einbeziehung unterschiedlicher Tonhöhen in den Stimuli vor. Für eine MDS-Untersuchung nach diesem Modell postulieren sie drei Szenarien, nachdenen sich die MDS-Struktur organisieren könnte:[16] Bei einer ausschließlichen Beurteilung der Klänge anhand der Klangquellen, würde sich im MDS-Ergebnis für jede Quelle ein Stimuli-Cluster bilden. Zweitens würde es bei einer ähnlichen Kategorisierung nach Klangquellen, in der jedoch nicht von Tonhöhenunterschieden abgesehen werden kann, zu einer Trennung der Quellen kommen. Es würde hierbei entlang einer zusätzlichen Dimension zu einer

13 Miller & Carterette (1975), Erickson (2003), Handel & Erickson (2004).
14 Marozeau et al. (2003).
15 Handel & Erickson (2004). S. 587-592.
16 ebd. S. 590.

Aufteilung nach der Tonhöhe kommen. Im dritten Szenario wird dagegen keine Kategorisierung nach Klangquellen angenommen, wobei hier ausschließlich physikalisch-akustische Eigenschaften, denen auch die Tonhöhe angehört, die MDS-Struktur bestimmen.

In ihrem Versuch verwendeten Handel & Erickson als Versuchs-Stimuli Instrumentalklänge aus der *McGill University Master Samples* (MUMS) Instrumentalklänge-Bibliothek. Als Instrumente wählten sie Trompete in C, Englischhorn und Klarinette in B, welche jeweils in fünf unterschiedlichen Tonhöhen vorlagen (g, $c^1$, $g^1$, $c^2$ und $g^2$). Die Klangqualität entsprach einer Abtastfrequenz von 44,1 kHz und einer Dynamikauflösung von 16-Bit. Die ursprünglich stereophonen Klänge wurden in monophone Klänge umgewandelt. Alle Stimuli wurden auf eine einheitliche Dauer von 2,0 sec editiert, wobei der Einschwingvorgang hiervon unberührt blieb. Um unerwünschte Klickgeräusche am harten Schnitt nach 2 sec zu vermeiden, wurde mittels einer *Spline*-Glättung ein künstlicher Ausschwingvorgang simuliert. Über die Dauer des Ausschwingvorgangs geben die Autoren keine Auskunft. Unklarheit herrscht zudem über die von ihnen vorgenommene Standardisierung der Lautstärke. Es kann lediglich aus einem ähnlichen Versuch mit Singstimmen[17] angenommen werden, daß sich die Standardisierung nur auf eine Normalisierung nach dem Effektivwert (*root-mean-square)* der logarithmierten Pegelgröße beschränkte.

Aus den insgesamt 15 Stimuli (3 Instrumente x 5 Tonhöhen) wurden eine Anzahl von 105 zu bewertender Paare von Stimuli gebildet. Diese Paare bildeten alle möglichen Kombinationen der Stimuli, ausgenommen der Paarungen von gleichen Stimuli und jenen Paarungen mit vertauschter Reihenfolge. Zwischen den Stimuli eines Paars gab es eine Pause von 500 msec. Jedes Paar wurde ein weiteres Mal wiederholt; in der Wiederholung mit vertauschter Stimuli-Reihenfolge. Vor dem Versuch wurden zum Kennenlernen alle Stimuli vorab präsentiert und zudem Beispielpaare vorgestellt.

Die Bewertung der Paare erfolgte mit Hilfe einer acht-stufigen, numerierten Ratingskala. An den Extremen gab es verbale Ankerbegriffe: *very similar* für ‚1‘ und *very different* für ‚8‘. Die Versuchsanweisung an die Versuchspersonen (VP) lautete in etwa:

> „[J]udge the difference in the quality of the two sounds, that part of a sound that distinguishes one instrument at two pitches or two instruments from one another. […] [D]isregard any differences in pitch, loudness, vibrato, and duration and […] judge only the difference in the way they soun[d].“[18]

Am Versuch nahmen 20 VP teil, welche entweder Studenten oder Graduierte der University of Tennessee waren. Keine der VP besaß umfangreiche musikalische Erfahrung bzw. Vorbildung. Des weiteren verfügte keine über ein eingeschränktes Hörvermögen. Die VP nahmen in einer schallgedämmten Kabine einzeln am Versuch teil. Die Wiedergabe der Stimuli erfolgte binaural (in Mono) über einen Kopfhörer. Bewertungen wurden direkt in einen Computer eingegeben. Jede VP erhielt eine randomisierte Reihenfolge von Paaren. Letztlich bleibt zu erwähnen, daß die VP an zwei

17 Erickson (2001).
18 Handel & Erickson (2004). S. 593.

unterschiedlichen Experimenten teilnahmen, wovon der hier geschilderte Versuch einen darstellte. Eine Hälfte der VP absolvierte den MDS-Versuch vor dem anderen, die andere Hälfte vice versa.

Als MDS-Ergebnis erhalten Handel & Erickson eine dreidimensionale Lösung und damit eine räumliche, geometrische Verteilung aller Stimuli. In Dimension 1 bildet sich bei ihnen ein starker Tonhöheneinfluß ab, welcher Anzeichen dafür liefert, daß die Verwendung unterschiedlicher Tonhöhen ohne einen Einfluß der Tonhöhenunterschiede nicht möglich ist. Darüber hinaus dominiert dieser angenommene Tonhöheneinfluß als strukturgebende Größe die geometrische Konstellation, in der sich die einzelnen Instrumente entlang Dimension 1 entsprechend der Tonhöhe erstrecken. Dimension 2 korreliert stark mit der Vergleichsgröße *spectral centroid*. In dieser Dimension zeichnet sich entlang Dimension 1 eine nicht ganz einwandfreie Trennung der Instrumente ab. Dimension 3 ist weitaus weniger aussagekräftig für das Gesamtergebnis und läßt sich, den Autoren nach, eventuell auf den Umfang der Variabilität des Vibratos in den Klängen zurückführen.

Der strukturgebende Einfluß der Tonhöhe, zusammen mit der nicht einwandfreien Trennung der Instrumente, deutet auf das Vorliegen einer MDS-Struktur, welche dem dritten Szenario von Handel & Erickson entspricht. Somit kommt es nicht zu der von ihnen erwarteten Kategorisierung nach ‚Quell-Klangfarben', welche die Einbindung mehrerer Tonhöhen begünstigen sollte. Sie folgern hieraus, daß paarweise Bewertungen von Klängen einen unabhängigen Vergleich der Ähnlichkeit von ‚Quell-Klangfarben' möglicherweise nicht zulassen.[19] Da nur zwei Stimuli miteinander verglichen werden, könnte es sein, daß mehrere Tonhöhen trotzdem nicht ausreichen, um die nötige Information für die Extrahierung von ‚Quell-Klangfarben' zu liefern. Handel & Erickson sind abschließend der Meinung, daß MDS-Untersuchungen mit Ähnlichkeitsbewertungen von Klangpaaren sich für die Untersuchung der ‚Quell-Klangfarbe' nicht eignen.

## 1.3 - Zielsetzung dieser Arbeit

Aus dem Versuchsplan von Handel & Erickson (2004) können einige Aspekte, unabhängig von den zuletzt genannten Gründen, eine Kategorisierung nach der ‚Quell-Klangfarbe' verhindert haben. Das Fazit von Handel & Erickson, wonach sich MDS-Untersuchungen für das Nachweisen von Kategorisierung nach Klangquellen möglicherweise nicht eignen, wird diesbezüglich in Frage gestellt. Ein Ziel dieser Arbeit ist es, anhand eines vergleichbaren Versuchs zu überprüfen, ob sich das Ergebnis von Handel & Erickson (2004) reproduzieren läßt, oder ob sich nicht doch eine Kategorisierung nach Klangquellen als Ergebnis bewahrheiten könnte. Hierbei sollen bezüglich der erwähnten Aspekte Änderungen vorgenommen werden.

In dieser Arbeit wird weiterhin einem explorativen Ansatz nachgegangen, ohne vorab Annahmen bzw. Hypothesen zu stellen. Durch einen geeigneten Versuchsplan für die Datenerhebung, sowie einer Standardisierung der Versuchs-Stimuli, soll ein Rahmen gegeben werden, aus dem sich nach Durchführung einer MDS das Ergebnis frei erkunden und interpretieren läßt. Trotz des Verzichts

19 Handel & Erickson (2004). S. 598-599.

auf a priori postulierte Hypothesen zugunsten des explorativen Ansatzes, können dennoch konkret anzustrebende Ziele für diese Arbeit formuliert werden:

a) Als erstes soll durch eine direkte Anlehnung an den Versuchsplan von Handel & Erickson (2004) eine gute Vergleichbarkeit ermöglicht werden. So kann die Reproduzierbarkeit ihrer Ergebnisse überprüft werden.

b) Zugleich soll jedoch in dieser Arbeit auch keine bloße Kopie des Versuchs von Handel & Erickson (2004) erfolgen, sondern auch Raum für einige Änderungen gelassen werden. Diese werden im folgenden dargelegt:

   i) Gegenüber Handel & Erickson (2004) soll eine Auswahl differenzierterer Tonhöhen getroffen werden. Diese sollen bezüglich der aus ihnen resultierenden Intervalle eine bessere qualitative Unterscheidbarkeit begünstigen.
   ii) Neben der Vergleichbarkeit mit Handel & Erickson (2004) sollen zugunsten neuer Erkenntnisse auch abweichende Instrumente eingebunden werden. Als Gemeinsamkeit sollen sich die Stimuli jedoch ausschließlich auf Klänge von Blasinstrumenten beschränken. Als weitere Neuerung soll zusätzlich ein Instrument in zwei unterschiedlichen Klangvarianten eingesetzt werden.
   iii) Änderungen gegenüber der Stimuli-Standardisierung sollen zugunsten bewertungsfreundlicher, sowie gehörgerechter Aspekte vorgenommen werden. Die Stimuli-Dauer von 2 sec bei Handel & Erickson (2004) wird diesbezüglich als zu kurz erachtet und soll verlängert werden. Zudem soll die Standardisierung der Lautstärke sich an eine dem menschlichen Hörempfinden angepaßte Größe richten. Hierfür ist die psychoakustisch-skalierte Größe Verhältnislautheit in Sone vorgesehen.
   iv) Abschließend soll außerdem eine gehörgerechte Abhörsituation für die Bewertung ermöglicht werden. Hierfür wird sich für die Aufzeichnung der Stimuli zugunsten der Kunstkopf-Stereophonie entschieden, welche bei einer binauralen Wiedergabe mit Kopfhörer eine dem Menschen gehörgerechte Wahrnehmung ermöglicht. Um Raumeinflüsse zu minimieren, soll des weiteren die Stimuli-Aufzeichnung in einer reflexionsarmen Umgebung durchgeführt werden.

c) Wie in der MDS-Forschung üblich, sollen auch hier quantifizierbare Vergleichsgrößen zur Deutung der dimensionalen Struktur einbezogen werden. Neben den gängigen Größen sollen auch einige aus der Psychoakustik stammende Vergleichsgrößen berücksichtigt werden. Es soll damit vor allem untersucht werden, inwiefern sie sich zur Beschreibung der Klangfarbe eignen.

d) Neben der Einbeziehung der quantitativen physikalisch-akustischen bzw. psychoakustischen Vergleichsgrößen, sollen zudem auch Ergebnisse aus der qualitativen Auswertung physikalisch-akustischer Analysen aufgenommen werden.

Mit dieser Zielsetzung vor Augen kann der Versuch konzipiert und durchgeführt werden, bevor im Anschluß an die Datenerhebung die MDS folgt. In Kapitel 2 wird an erster Stelle als wichtigster Bestandteil des Versuchs die Frage der Natur und Zusammensetzung der Stimuli, sowie ihrer Produktion geklärt. Des weiteren wird dort bereits eine umfassende qualitative Betrachtung der physikalisch-akustischen Eigenschaften der Stimuli vorgenommen. Die Konzeption und Durchführung des Versuchs schließt sich in Kapitel 3 an. Es folgen in Kapitel 4 die ersten Schritte nach der Datenerhebung mit der Auswertung der Daten, sowie am Ende die endgültige Bestimmung eines Ergebnisses der MDS. Eine ausführliche Präsentation des MDS-Ergebnisses, inklusive der quantitativen und qualitativen Deutung ihrer Struktur, erfolgt in Kapitel 5. Abschließend wird in Kapitel 6 das Ergebnis mit Blick auf Handel & Erickson (2004), sowie weiterer relevanter Studien diskutiert und ein Ausblick auf zukünftig einzuschlagende Wege in der Klangfarbenforschung geboten.

# 2 - AUSWAHL DER INSTRUMENTE & TONHÖHEN

Zur Überprüfung der Ergebnisse von Handel & Erickson (2004) ist es nötig, zu ihrem Versuch eine gewisse Vergleichbarkeit, durch die Beibehaltung bestimmter Voraussetzungen, zu gewährleisten. Gleichzeitig soll jedoch auch die Möglichkeit genutzt werden, alternative Ansätze zu verfolgen bzw. eine in manchen Punkten anders realisierte Standardisierung vorzunehmen.

Der Versuch von Handel & Erickson (2004) beinhaltet 15 Stimuli, mit 3 Instrumenten x 5 Tonhöhen. Dies macht eine Bewertung von 105 Stimuli-Paaren notwendig. Für den Umfang dieser Magisterarbeit wird der Versuchsaufwand auf die Bewertung von 66 Paaren (entsprechend einer Anzahl von 12 Stimuli) begrenzt. Die 12 Stimuli setzen sich aus 4 Instrumenten bzw. Instrumental-variationen x 3 Tonhöhen zusammen.

## 2.1 - Auswahl der Instrumente

Die Instrumente von Handel & Erickson (2004) lassen sich in zwei Gruppen einteilen. Darin bilden die Klarinette in B und das Englischhorn eine Gruppe von Holzblasinstrumenten, während die Trompete in C den einzigen Vertreter der Blechblasfamilie darstellt. Um die Vergleichbarkeit zu Handel & Erickson (2004) zu gewährleisten, wird die B-Klarinette[20] als gemeinsames Instrument beider Versuche ausgewählt. Die Klarinette zeichnet sich dadurch aus, allein durch ihren großen Tonumfang, bereits über eine sehr breite Klangfarbenvielfalt zu verfügen. Neben der gebotenen Vergleichbarkeit zu Handel & Erickson (2004) handelt sich es also um ein besonders für die Untersuchung von Instrumentalklangfarben auf unterschiedlichen Tonhöhen geeignetes Instrument. Aufgrund der begrenzten Anzahl der Instrumente werden für die restlichen Instrumente von Handel & Erickson (2004) abweichende Instrumente gewählt. Um jedoch auch hier eine Vergleichbarkeit zu gewähren, werden ausschließlich Blasinstrumente ausgewählt: Baßflöte und Horn.

Der Baßflöte[21] wird gegenüber der gebräuchlichen großen Querflöte in C der Vorzug gegeben, da sich ihr Tonbereich eher mit dem der B-Klarinette, jedoch insbesondere mit dem des Horns vergleichen läßt. Es wird erwartet, daß die Baßflöte, trotz ihres tieferen Klangs, für Querflöten typische Klangeigenschaften aufweist. Mit dem Horn[22] wird analog zu Handel & Erickson (2004) ein Vertreter der Blechblasfamilie gewählt. Im Vergleich zu allen Blechblasinstrumenten weist das Horn klanglich jedoch die größte Nähe zu den Holzbläsern auf.

Zu den drei Instrumenten soll zusätzlich noch eine Instrumentalvariation hinzukommen. Als Instrumentalvariation dient eine klangliche Abwandlung eines der verwendeten Instrumente. Variiert wird hierbei das Horn, mit Klängen *ohne* und *mit* Dämpfer. Diese zwei Instrumentalvariationen des

20 Es ist anzumerken, daß als Besonderheit am Rohrblatt des Instruments zusätzlich eine dünne Aluminiumfolie angebracht wurde.

21 Bei der Baßflöte handelt es sich um das Baßinstrument der Querflöten-Familie. In diesem Fall handelte es sich ferner um ein Vierteltoninstrument (mit zusätzlicher Klappe für die Erzeugung von Viertelton-Intervallen).

22 Beim Horn handelt es sich um ein Deutsches Doppelhorn, mit durch Ventil umschaltbarer Grundstimmung in F oder B. Mit Ausnahme der Tonhöhen $c^1$ und $f^1$ wurden alle weiteren Tonhöhen in der B-Grundstimmung gespielt.

Horns beruhen auf einer identischen Klangerzeugung, jedoch mit unterschiedlichen Schwingungs- und Abstrahleigenschaften im Bereich des Schalltrichters. Hiermit soll ermöglicht werden, Unterschiede zwischen Instrumenten und Unterschiede zwischen Klangvariationen eines Instruments im MDS-Ergebnis vergleichen zu können.

## 2.2 - Auswahl der Tonhöhen

Wie bereits erwähnt sollen für den Versuch selbst nur 3 Tonhöhen verwendet werden, da mit der Hinzunahme jeder weiteren Tonhöhe die Anzahl der Stimuli und damit vor allem die Anzahl zu bewertender Klangpaare deutlich zunehmen würde. Bei der Stimuli-Aufzeichnung selbst werden allerdings 5 Tonhöhen aufgezeichnet. Auf diese Weise soll ein besseres Verständnis über die tonhöhenabhängige Entwicklung der physikalisch-akustischen Eigenschaften erhalten werden. Sollten sich nach Auswertung der Aufnahmen Probleme bzw. Inkompatibilitäten der Klänge ergeben, bietet die größere Anzahl von Tonhöhen außerdem Ausweichmöglichkeiten. Die Tonhöhen sollen sich über einen Tonumfang erstrecken, in dem sich die Instrumente sinnvoll untereinander vergleichen lassen. Aufgrund des eher tieferfrequent angesiedelten Horns können bei der Klarinette nur die tieferen Register berücksichtigt werden. Der Tonumfang umfaßte einundhalb Oktaven, was für drei Versuchs-Stimuli geeignet scheint.

Nach der Zielsetzung soll eine differenziertere Tonhöhenauswahl getroffen werden als bei Handel & Erickson (2004). Ihre Tonhöhen (c, g, $c^1$, $g^1$ und $c^2$) ergaben sich aus einer bloßen Oktavierung der zwei Tonhöhen c und g. Mit der überwiegenden Beschränkung auf Intervalle einer Oktave bzw. ihrer Vielfachen könnte die ihnen eigene qualitative Nähe bzw. das Phänomen der Oktavengleichheit Ähnlichkeitsbewertungen der Klangfarbe zusätzlich beeinflussen. Des weiteren birgt die ausschließliche Verwendung der Tonhöhen c und g (also das Intervall einer Quinte) weitere Nachteile, da die Tonhöhen dieses Intervalls zueinander naturgemäß einen hohen Verwandtschaftsgrad aufweisen. Die Bezeichnung ‚reines Intervall' für Oktave und Quinte zeugt davon. Zudem lassen sich vorhandene Formantbereiche beim Auftreten von immer gleichen Frequenzlagen von Teiltönen über den einzelnen Stimuli-Tonhöhen schlechter abbilden. Es sollen für diese Arbeit daher Tonhöhen gewählt werden, welche einen möglichst geringen Verschmelzungsgrad aufweisen und sich somit besser voneinander unterscheiden lassen.

Anhand der vorangegangenen Überlegungen werden schließlich die Stimuli in den folgenden 5 Tonhöhen aufgezeichnet: g, $c^1$, $f^1$, $h^1$ und $d^2$. Auf die Verwendung der höchsten Tonhöhe $d^2$ wurde später jedoch ganz verzichtet, da diese nicht mehr sauber vom Horn gespielt werden konnte. So kommen am Ende die Tonhöhen g, $c^1$ und $h^1$ als Versuchs-Stimuli zum Einsatz, während die Tonhöhe $f^1$ als zusätzliche Tonhöhe (nachfolgend ‚Zusatz-Stimulus' genannt) für die qualitative Auswertung der physikalisch-akustischen Analysen berücksichtigt wird.

## 2.3 - Die Stimuli und ihre Standardisierung

Entsprechend der Zielsetzung dieser Arbeit werden die zu verwendenden Versuchs-Stimuli in bestimmten Punkten einer Standardisierung unterzogen, um den Einfluß nicht erwünschter unabhängiger Variablen zu minimieren. Darüber hinaus soll die Standardisierung eine dem menschlichen Gehör gerechte Wiedergabe über Kopfhörer ermöglichen. Daher wird entschieden, die Stimuli selbst aufzuzeichnen und zu produzieren. In dieser Produktion können so Standardisierungen hinsichtlich ausgewählter Aspekte im Vorhinein realisiert werden.

### 2.3.1 - Aufnahme der Stimuli

Im Alltag nimmt der Mensch seine akustische Umwelt stets binaural wahr. Auch bei Konzertbesuchen oder beim Hören von stereophonen Aufnahmen wird binaural wahrgenommen. In der Vergangenheit wurden allerdings viele MDS-Versuche nur mit der Wiedergabe monophoner Stimuli durchgeführt. Im subjektiven Hörvergleich büßt die Reduzierung auf Mono hierbei jedoch eine gewisse ‚Lebhaftigkeit' bzw. Natürlichkeit im Klangeindruck ein. Obwohl diese Charakteristiken durchaus durch raumakustische Faktoren oder direktionale Ortung von Schallquellen begünstigt werden, ist diese empfundene ‚Räumlichkeit' sicherlich ebenfalls Bestandteil unseres Verständnisses von Klangfarbe. Daher wird für die Aufzeichnung der Stimuli die Kunstkopf-Stereophonie gewählt, welche der menschlichen binauralen Hörwahrnehmung am besten entspricht. Ferner soll durch die Wahl eines geeigneten Aufnahmeraums die Möglichkeit gegeben sein, raumakustische Faktoren stark einzuschränken.

In der konkreten Umsetzung wurden die Stimuli mit dem digitalen Kunstkopf-Meßmikrophonsystem *HMS III* der Firma HEAD Acoustics aufgezeichnet. Kunstkopfaufnahmen berücksichtigen als direkte Äquivalente zu der menschlichen Hörwahrnehmung die menschliche Kopfübertragungsfunktion, inklusive der Einwirkung der physiologischen Voraussetzungen der Ohrmuschel und des äußeren Gehörkanals.[23] Die Aufzeichnung der Stimuli erfolgte im Zusammenspiel mit dem HEAD Acoustics *Audio Recorder*, welcher eine optimierte, verlustfreie Einbindung in das HEAD Acoustics Programm *ArtemiS* sicherstellte. *ArtemiS* stellt eine sehr umfangreiche Analyse-Umgebung dar, mit welcher einige akustische Analysen dieser Arbeit durchgeführt werden konnten.

Als Aufnahmeraum kam der reflexionsarme Raum (RAR) des Instituts für Technische Akustik an der Technischen Universität Berlin zum Einsatz. Dieser Raum ist voll-reflektionsfrei ausgestattet (mit hoher Absorption an allen 6 raumbegrenzenden Flächen) und erfüllt in der Regel akustische Freifeldbedingungen ab einer Frequenz von 36 Hz aufwärts.[24] Freifeldbedingungen gewährleisten

23 Um eine 100-prozentige gehörgerechte Wiedergabe von Aufnahmen mit dem HEAD Acoustics Kunstkopf gewährleisten zu können, ist ausgangsseitig zudem eine spezifische Verstärker-Kopfhörer-Kombination nötig, welche eine Entzerrung der Kopf- und Ohrmuschelübertragungsfunktion vornimmt. Aus praktischen Gründen konnte diese nicht eingesetzt werden, wodurch sich keine 1:1-Nachbildungen der Aufnahmen realisieren ließen. Dieser Nachteil beschränkt sich jedoch nur auf Unterschiede in der Frequenz-Übertragungsfunktion, welche sich auf beide Hörkanäle identisch auswirken und ferner alle Stimuli gleichermaßen beeinflussen. Somit war für die ohnehin relativen Bewertungen kein systematischer Fehler zu erwarten. Interaurale Unterschiede blieben hiervon unberührt, wodurch die Qualität der stereophonen Übertragung in keinster Weise beeinträchtigt wurde.

24 Im Vergleich dazu beträgt die Grundtonfrequenz der tiefsten aufgenommenen Tonhöhe g ca. 197 Hz.

in der Regel eine Aufnahme der Stimuli, bei der nur die akustischen Übertragungseigenschaften des Instruments, ohne Einfluß der Übertragungseigenschaften des Raums (Raumresonanzen, -reflexionen und Nachhall), in Erscheinung treten. Dies ist sowohl für die physikalisch-akustische Analyse als auch für die Abhörsituation bei der Bewertung der Instrumentalklänge von Vorteil.

Nichtsdestotrotz stellen Aufnahmen aus einer reflexionsarmen Umgebung eine im Alltag nicht bekannte Hörsituation dar, da man in der Regel gerade daran gewöhnt ist, eine gewisse akustische ‚Räumlichkeit' in einem Klangeindruck vorzufinden. Andererseits ist man ebenso daran gewöhnt, gleiche Instrumente unter verschiedenen raumakustischen Umständen zu hören, so daß angenommen werde kann, daß dies auch für den Extremfall des RAR zutreffen würde. Um die Verwendbarkeit der im RAR aufgenommenen Stimuli zu überprüfen, wurden vorab Testaufnahmen mit anschließenden subjektiven Bewertungen durch mehrere Personen durchgeführt. Dabei stellte sich heraus, daß die Aufnahmen eine gewisse ‚Natürlichkeit' des Höreindrucks bewahren, wodurch der Klangeindruck im RAR vergleichbar mit denen unter realen akustischen Bedingungen ist.

Um einen der Realität entsprechenden Klangeindruck der Instrumente zu gewährleisten, war es zudem notwendig, frequenzabhängige Abstrahlcharakteristiken der Instrumente zu berücksichtigten, da im akustischen Freifeld von der Instrument-Mikrophon-Achse stärker abweichende Abstrahlrichtungen nicht mit aufgezeichnet werden. Zu Rate gezogen wurden hierzu Erkenntnisse über die Abstrahlcharakteristik von Orchesterinstrumenten von Meyer. Wichtig war es, eine einzige Abstrahlrichtung des jeweiligen Instruments zu ermitteln, unter der alle Frequenzbereiche annähernd gleich stark abgestrahlt werden.

Für die Klarinette wird laut Meyer seitlich vom Musiker bei einem Winkel von 50° oberhalb der Längsachse des Instruments (Mundstück-Trichter-Achse) (siehe Abb. 2-1) die ausgewogenste breitbandige Abstrahlung erreicht.[25] Jedoch sind selbst in diesem Fall über 10 dB große Differenzen zwischen unterschiedlichen Frequenzen zu erwarten, da einzelne Frequenzbereiche sehr gerichtet abgestrahlt werden und somit in einer Abstrahlrichtung nicht alle Bereiche gleich stark vertreten sind. Ein Vergleich der Aufnahmen aus dem RAR mit Aufnahmen aus reflektierenden Räumen lieferte sowohl für die Spektralanalysen als auch für die subjektiven Höreindrücke keine grundlegende Verfremdung des Klarinettenklangs.

25 Meyer (1999). S. 120-121.

Abb. 2-1

Das Horn strahlt alle Frequenzen nur gleich stark in Richtung der Trichterachse ab.[26] In der gewöhnlichen Hör- bzw. Aufführungssituation wird der Schalltrichter schräg nach hinten-rechts gehalten. Durch Reflexion an Rückwänden werden gewöhnlich alle Frequenzbereiche nach vorne, d.h. in Richtung des Publikums bzw. der Mikrophone, abgestrahlt. Im RAR gilt dies nicht, womit das Horn direkt aus der Trichterachse aufgenommen werden mußte (siehe Abb. 2-2).

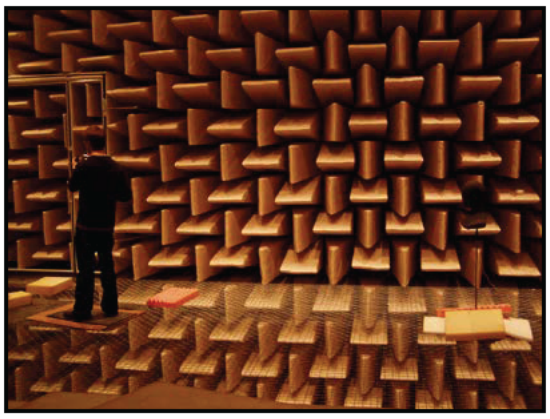

Abb. 2-2

Zur Abstrahlcharakteristik von Querflöten liefert Meyer nur Informationen für die große C-Flöte.[27] Die tonalen Frequenzanteile (harmonische Obertonreihe) werden ausgewogen unter einem Winkel von 20-30° von einer idealen Linie senkrecht zur Instrumentenachse abgestrahlt, d.h. einer Linie in Blickrichtung des Musikers. Da für die Baßflöte, bis auf die größeren Dimensionen und ein in U-Form gebogenes Ende zum Mundstück hin, das akustische Wirk- und Abstrahlprinzip weitestgehend vergleichbar mit der großen Flöte ist, wurde die sie in Richtung des oben genannten Winkels aufgenommen (siehe Abb. 2-3). Diese Richtung wurde auch von der Flötistin als die ihr bekannte Hauptabstrahlrichtung bestätigt. Gegenüber den tonalen Frequenzanteilen werden die für die Flöte charakteristischen Geräuschanteile kugelförmig abgestrahlt, so daß hierfür keine spezielle Vorzugsrichtung zusätzlich zu berücksichtigen war.

26 Meyer (1999). S. 113-115.
27 ebd. S. 116-118.

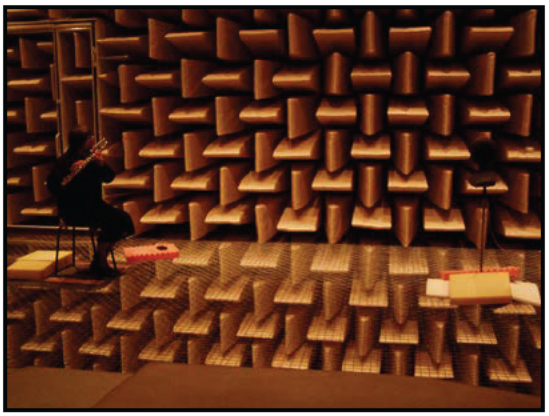

Abb. 2-3

Die Aufnahmen wurden, wie in Abb. 2-1 bis 2-3 zu sehen, mit einem Mikrophonabstand von ca. 3 m durchgeführt. Die Aufnahmen wurden mit im HEAD Acoustics *Audio Recorder* eingestellter Freifeld-Entzerrung durchgeführt, welche einen Abstand zur Schallquelle von 3 m voraussetzt.[28] Die Musiker erhielten vor den Aufnahmen die Gelegenheit sich im RAR einzuspielen, um sich an die ungewohnte Akustik gewöhnen zu können. Vor der Aufnahme einer Tonhöhe wurden die Instrumente mit einem Stimmgerät auf den Zielton eingestimmt (Grundstimmung: $a^1$ = 443 Hz). Die Musiker wurden aufgefordert, alle Tonhöhen mit einer konstanten Artikulation und Dynamik und ohne Vibrato zu spielen. Darüber hinaus sollten die Töne gleichmäßig gespielt und länger als 5 sec gehalten werden. Außerdem wurde darauf geachtet, unter allen Instrumenten eine ähnliche Artikulationsart zu erreichen: generell ein weicher Blasansatz und eine klanglich vergleichbare Dynamik. Jede Tonhöhe wurde fortlaufend mindestens fünfmal aufgezeichnet, um später daraus einen, für den Versuch geeigneten, Ton auswählen zu können.

### 2.3.2 - Standardisierung der Stimuli

Nach der Auswahl der in Frage kommenden Klänge wurde versucht die Klänge in ihrer Dauer und Lautstärke zu standardisieren.

*a) Standardisierung der Stimuli-Dauer*

Die bei Handel & Erickson (2004) verwendete einheitliche Stimuli-Länge von 2 sec wurde für diese Arbeit als zu kurz erachtet, da diese Dauer eventuell nicht ausreicht, um einen Eindruck von der Klangfarbe eines Stimulus zu erhalten oder den Vergleich zweier Stimuli zu ermöglichen. Um diesem Problem begegnen zu können, wurde eine einheitliche Dauer von 5 sec gewählt. Die Verwendung von kurzen Stimuli-Dauern birgt allerdings auch Vorteile für die Versuchsdurchführung. Durch kürzere Stimuli-Dauern verkürzt sich nämlich gleichzeitig auch der zeitliche Versuchsaufwand. Bei Versuchen mit einer großen Anzahl von Stimuli, welche eine noch weitaus größere Anzahl von zu bewertenden Paaren nötig macht, kommt eine Verkürzung sehr zugunsten einer besseren und konsistenten Bewertungsfähigkeit und -leistung der VP. Es sei nebenbei angemerkt, daß

28 Aufgezeichnet wurde zwei-kanalig im HEAD Acoustics eigenen Audio-Datenformat. Die Aufnahmequalität entspricht einer Abstastrate von 44,1 kHz mit einer Dynamikauflösung von 16-Bit.

Kendall et al. den Einfluß der Stimuli-Dauer (mit zwei unterschiedlichen Dauern von 0,5 und 3 sec) in einem Identifikationstest für Instrumentalklänge untersuchten und keine Anhaltspunkte dafür finden, daß die Stimuli-Dauer die Identifikation signifikant beeinflußt.[29] Diese Feststellung läßt sich jedoch nicht direkt auf die Bewertung von Ähnlichkeiten anwenden, da die Aufgabenstellung in beiden Versuchen bereits zu unterschiedlich ist.

Alle Stimuli dieser Arbeit wurden samplegenau auf 5 sec begrenzt. Der natürliche Einschwingvorgang der Klänge, welcher für die Klangfarbenwahrnehmung eine wichtige Rolle spielen kann, blieb davon unberührt. Demnach mußte der Ausklingvorgang künstlich herbeigeführt werden. Dies wurde bekanntlich auch bei den Stimuli von Handel & Erickson (2004) vorgenommen. Sie berufen sich dabei auf Versuche von Saldanha & Corso, welche ergaben, daß der Ausklingvorgang unterschiedlicher Blasinstrumente praktisch nicht zur korrekten Identifikation der Instrumente beiträgt.[30] Der künstliche Ausklingvorgang wurde mit einer standardisierten exponentiellen Abklingfunktion mit einer einheitlichen Dauer von 60 msec umgesetzt. Diese Dauer wurde von mehreren Personen subjektiv als natürlich und den Klang nicht verfälschend bewertet.

*b) Standardisierung der Lautstärke*

Aus der Zielstellung ist bekannt, daß zur Standardisierung der Stimuli-Lautstärke die psychoakustische Größe der *Verhältnislautheit* in Sone genutzt werden sollte. Diese psychometrisch-skalierte Größe drückt die subjektiv empfundene Lautstärke aus. Die Berechnung und Anpassung der Verhältnislautheiten konnte mit *ArtemiS* realisiert werden. Hierzu wurden vorab Einzahlwerte für die Verhältnislautheit in Sone aller 12 Stimuli (sowie der 4 Zusatz-Stimuli) ermittelt. Diese sind zusammen mit den entsprechenden Schalldruckpegeln in dB in Tabelle 2-1 dargestellt.

29 Kendall et al. (1999). S. 336.

30 Saldanha & Corso (1964) ließen VP Hörstimuli in vorgegebene Instrumentalkategorien klassifizieren. Als Instrumente kamen unter anderem auch Klarinette, große Querflöte und Horn in drei Tonhöhen ($c^1$, $f^1$ und $a^1$) zum Einsatz. Während zeitliche Alterationen der stationären Phase und das Weglassen des Einschwingvorgangs eines Klangs zu signifikanten Unterschieden in der Instrumentenidentifikation beitrugen, führte das Weglassen des Ausklingvorgangs in keinem Fall zu einer besseren oder schlechteren Identifikation.

| Stimuli | Verhältnislautheit N in Sone | | Schalldruckpegel $L_{SPL}$ in dB | |
|---|---|---|---|---|
| Instrument & Tonhöhe | links | rechts | links | rechts |
| ***Versuchs-Stimuli*** | | | | |
| Baßflöte [g] | 14.90 | 14.90 | 72.05 | 72.39 |
| Baßflöte [$c^1$] | 14.50 | 15.30 | 73.85 | 74.81 |
| Baßflöte [$h^1$] | 15.10 | 14.60 | 72.04 | 72.00 |
| Klarinette [g] | 15.60 | 15.60 | 68.13 | 68.15 |
| Klarinette [$c^1$] | 15.60 | 15.80 | 66.78 | 67.10 |
| Klarinette [$h^1$] | 16.30 | 16.10 | 68.94 | 68.81 |
| Horn normal [g] | 15.00 | 15.10 | 70.42 | 70.38 |
| Horn normal [$c^1$] | 15.00 | 15.00 | 70.47 | 70.50 |
| Horn normal [$h^1$] | 14.90 | 14.90 | 73.26 | 73.20 |
| Horn mit Dämpfer [g] | 14.30 | 14.80 | 64.74 | 65.43 |
| Horn mit Dämpfer [$c^1$] | 15.20 | 15.40 | 68.39 | 68.30 |
| Horn mit Dämpfer [$h^1$] | 15.50 | 15.60 | 71.50 | 71.47 |
| ***Zusatz-Stimuli*** | | | | |
| Baßflöte [$f^1$] | 15.30 | 14.50 | 70.49 | 69.44 |
| Klarinette [$f^1$] | 15.90 | 15.90 | 66.35 | 66.43 |
| Horn normal [$f^1$] | 16.20 | 15.70 | 72.99 | 72.66 |
| Horn mit Dämpfer [$f^1$] | 15.20 | 15.40 | 70.53 | 70.42 |

Tabelle 2-1

Wie zu sehen ist, besitzt die Baßflöte eine deutlich geringere Verhältnislautheit als die anderen Instrumente. Um alle Lautheiten anzunähern, war es jedoch nicht sinnvoll, als Standardwert die Mitte des vorliegenden Umfangs zu wählen. Durch die geringere Aussteuerung und der damit verbundenen geringeren Dynamikauflösung (Bit-Breite), sowie des niedrigeren Signalrauschabstands, hätten die Stimuli der Baßflöte so an Klangqualität verloren. Stattdessen wurde ein Standardwert von 15 Sone gewählt, so daß die Flötenklänge in der Lautheit höchstens verdoppelt werden mußten. Als Resultat wurde ebenfalls das Grundrauschen in seiner Lautheit verdoppelt, so daß es im Vergleich zu den anderen Klängen einfacher wahrzunehmen war. Die Differenz des Grundrauschpegels von Baßflöte zu denen der Klarinette und des Horns mit Dämpfer betrug etwa 10 dB. Ein niedrigerer Standardwert als 15 Sone hätte in diesem Fall nicht geholfen, da es zu keiner Minderung des bestehenden Unterschieds im Grundrauschen beigetragen hätte. Spätestens in der Abhörsituation, bei entsprechender Verstärkung, wäre es weiterhin ungünstig in Erscheinung getreten. Als einzige Alternative wurde letztlich eine standardisierte Rauschunterdrückung um ca. 10 dB auf alle Baßflötenklänge angewendet.[31] Durch die Möglichkeit die (durch die Rauschunterdrückung) entfernten Klanganteile (getrennt vom entrauschten Flötensignal) akustisch wiederzugeben, konnte subjektiv sichergestellt werden, daß keine tonalen Klanganteile und Transienten aus den Flötenklängen entfernt wurden. Durch die Rauschunterdrückung wurden sehr geringfügig Blasgeräuschanteile bei allen Flötenklängen gleich stark entfernt. Nach dieser Nachbearbeitung der Baßflöten-Stimuli wurden alle Klänge bei einer annehmbaren Toleranz

31 Verwendet wurde Sonic Foundry *Noise Reduction* (Version 2.0). Dieses Programm erstellt aus einem Teil des isolierten Grundrauschens ein sogenanntes *Noise Print*. Das *Noise Print* wird als Grundlage für die Rauschunterdrückung genutzt um gezielt die im *Noise Print* enthaltenen Störanteile zu entfernen, und zwar bei einer minimalen Beeinflussung des Nutzsignals. Die Analyse des Audiosignals stützte sich auf eine *Fast Fourier Transformation* (FFT) mit 16384 Frequenzlinien (77 % Fensterüberlappung).

von ±1 Sone auf 15 Sone normalisiert.[32] Bereits die Unterschiede in der Lautheit zwischen linkem und rechtem Kanal bewegen sich in dieser Größenordnung (siehe Tabelle 2-1), womit auf eine noch kleinere Toleranz verzichtet werden konnte.

Es sei jedoch angemerkt, daß selbst die Verwendung einer psychometrischen Größe nur in ihrer Berechnung integrierte Charakteristiken berücksichtigt und somit unter Umständen andere bestehende, wahrnehmbare Unterschiede ignoriert werden. Die Verhältnislautheit stützt sich diesbezüglich ausschließlich auf physikalisch-akustische Informationen. Auf einer eher semantischen Ebene könnten für die empfundene Lautstärke von musikalisch relevanten Klängen jedoch gleichzeitig auch Unterschiede in den Dynamikstufen (‚musikalische Lautheit') eine Rolle spielen. Als überspitztes Beispiel könnten so Klänge in zwei grundlegend unterschiedlichen Dynamikstufen (*pp* und *ff*) vorliegen, welche sich jedoch problemlos in der Verhältnislautheit in Sone anpassen lassen. Es müßte in solch einem Fall hinterfragt werden, ob trotz der gleichen Werte in der Verhältnislautheit, diese Klänge subjektiv wirklich als gleich laut empfunden werden.

Für die aufgezeichneten Klänge lag kein so drastischer Unterschied in den Dynamikstufen vor, da eine ähnliche Dynamik über alle Tonhöhen bei der Aufnahme gefordert wurde. Zugunsten eines sauber gespielten Tons war es zum Teil jedoch nötig auf spieltechnisch leichter zu bewerkstelligende Dynamikstufen zurückzugreifen, so daß unter Umständen aus abweichenden Dynamikstufen empfundene Lautstärkeunterschiede vorhanden sind. Somit lassen sich die oben genannten Faktoren nicht ausschließen. Allerdings kann ihr Einfluß als weniger bedeutsam eingeschätzt werden. Eine bessere Alternative der Lautstärken-Standardisierung wäre eine von mehreren Personen subjektiv durchgeführte Anpassung, da die oben genannten Faktoren bezüglich der Dynamikstufen in dieser Anpassung der empfundenen Lautstärke dann berücksichtigt wären.[33]

## 2.4 - Instrumentenspezifische Charakteristika über die gewählten Tonhöhen

An dieser Stelle werden für jedes Instrument über alle Tonhöhen die instrumentenspezifischen Charakteristiken bezüglich ihrer Frequenzspektren und temporalen Verläufe besprochen. Als Referenz dienen hier erneut Erkenntnisse von Meyer, welche mit den physikalisch-akustischen Analysen der Stimuli dieser Arbeit verglichen werden. Diese Ergebnisse werden schließlich bei der qualitativen Analyse des MDS-Ergebnisses einbezogen.

### 2.4.1 - Klarinette

Die Klarinette ist als eines der in ihrer Klangfarbe vielseitigsten Instrumente bekannt. Maßgeblich kommt diese Wandlungsfähigkeit des Klarinettenklangs entlang des Tonraums zur Geltung. In der Regel spricht man bei der Klarinette von drei im Klang charakteristisch einteilbaren Bereichen, den sogenannten Registern. In ihrer tiefsten Tonlage (von etwa e bis $c^1$) befindet sich das sogenannte

32 Alle ermittelten Verhältnislautheiten N in Sone wurden nach ISO532-Norm für Lautheit im Freifeld unter Verwendung von FFT-Analysedaten bestimmt. Die FFT wurde stets mit 4096 Frequenzlinien, Hanning-Fensterung und 50 % Fensterüberlappung durchgeführt.

33 vgl. Lakatos (2000) und Marozeau et al. (2003).

*chalumeau*-Register, welches sich durch einen hohlen, dunklen und sonoren Klang auszeichnet. Im mittleren Register (von etwa $d^1$ bis $c^2$) besitzt der Klang eine schwache und trübe Prägung. Das hohe *clairon*-Register (von etwa $d^2$ bis $c^3$) wird charakterisiert durch einen klaren und „silbernen" Klang.[34]

Der Tonraum der in dieser Arbeit verwendeten Stimuli (g bis $h^1$) beschränkt sich bei der B-Klarinette daher auf die beiden unteren Register. Es bleibt zu erwähnen, daß die Transformation der charakteristischen Klangfarbe von einem Register ins andere weitestgehend fließend erfolgt und somit die oben angegebenen Grenzen nur als Richtwerte zu verstehen sind, die zudem vom Instrument und der Anblastechnik abhängen. Rimsky-Korsakov meint zur Existenz unterschiedlicher Register bei Holzblasinstrumenten und ihrer Beziehung zueinander:

> „It is difficult to define the exact limits of each register; adjacent registers also blend together and the passage from one to another is scarcely noticeable. But when the instrument jumps from one register to another the difference on power and quality of tone is very striking."[35]

Zudem kommt, wie auch in der weiteren Betrachtung der Spektren zu sehen sein wird, daß sich bei der Klarinette die Spektralcharakteristik zwischen dem tiefsten und höchsten Register grundlegend unterscheidet. Das mittlere Register besitzt hierbei die Rolle eines Bereichs, in dem eine allmähliche Wandlung zwischen diesen beiden Extremen vollzogen wird. Die beiden Stimuli-Tonhöhen g und $c^1$ befinden sich demnach im tiefsten Register, während die höchste Stimuli-Tonhöhe $h^1$ sich im oberen Bereich des mittleren Registers befindet, d.h. eine Wandlung hin zur charakteristischen Klangfarbe des oberen Registers zu erwarten ist. Diese interne Gruppierung der Klarinetten-Stimuli wird im folgenden bei der Interpretation der Ergebnisse der MDS eventuell zu berücksichtigen sein.

Bezüglich der Spektralcharakteristik der Klarinette erwähnt Meyer[36] für das tiefste Register die Besonderheit, daß eine sehr ausgeprägte Dominanz der ungeradzahligen Teiltöne vorliegt. Besonders die ersten geradzahligen Teiltöne (die ersten beiden Oktavtöne) sind gegenüber der Ausprägung der ungeradzahligen sehr schwach ausgeprägt. Hieraus erklärt sich auch die ‚dunkle' Klangfarbe. Wie in Abb. 2-4 und 2-5 zu sehen, liegt diese auch bei den aufgenommenen Klarinetten-Stimuli vor.[37] Auch sieht man hier, wie bei dem Tonhöhenunterschied einer Quarte die Ausprägung der relativ starken ungeradzahligen Teiltöne bereits bei $c^1$ weniger stark ausgeprägt ist.

34 vgl. Rimsky-Korsakov (1964). S. 16-17, sowie Meyer (1999). S. 65-66.
35 Rimsky-Korsakov (1964). S. 14.
36 Meyer (1999). S. 65-68.
37 Alle Spektren wurden mittels FFT mit 32768 Frequenzlinien erstellt, entsprechend einer Frequenzauflösung von 1,35 Hz. Die Berechnung und graphische Darstellung erfolgte mit dem Programm Mathworks *MATLAB*, Version 6.5 R13.

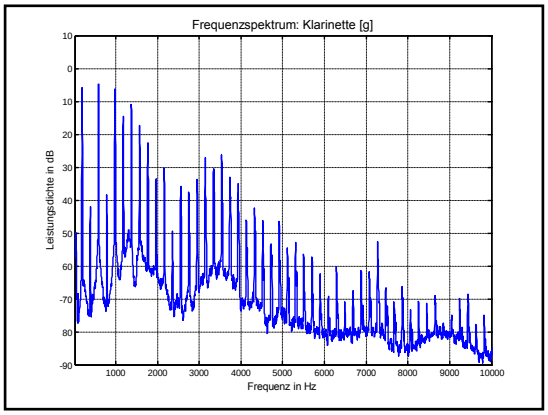


Abb. 2-4

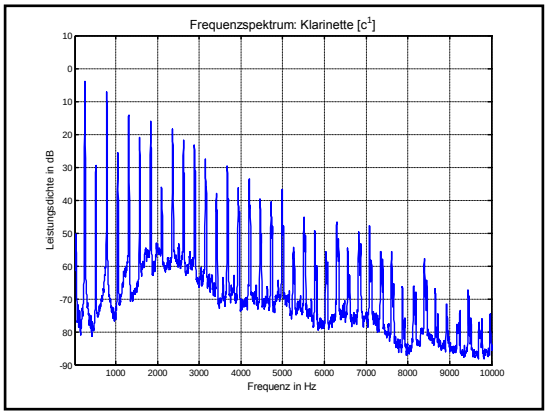


Abb. 2-5

Erwähnenswert ist bei der Tonhöhe g auch die deutliche Ausprägung eines Formantbereichs zwischen 3000 und 4000 Hz. Laut Meyer entspricht dieser der Vokalfarbe ‚i' und vermittelt „Glanz und Leuchtkraft"[38]. Die starke Ausprägung dieses Formanten in dieser Tonhöhe verwundert insoweit, daß Meyer ihn als Charakteristik des oberen Registers nennt.

Für das mittlere Register bzw. bei der Wandlung der Klangfarbe hin zum oberen Register tritt mit zunehmender Frequenz eine Abnahme der Differenz der Amplituden der ungeradzahligen und geradzahligen Teiltöne ein. Eine schwache Ausprägung des zweiten Teiltons (Oktavton) bleibt jedoch als prägendes Charakteristikum dieses Registers erhalten. Ab dem vierten Teilton stellt sich eine ausgewogene Beziehung zwischen allen Teiltönen ein. Es bildet sich für diesen Bereich eine stetig in der Amplitude abnehmende harmonische Obertonreihe. Diese Entwicklung läßt sich für den Zusatz-Stimulus $f^1$, wie in Abb. 2-6 zu sehen, bereits ansatzweise erkennen. Im unteren Bereich des mittleren Registers befindlich, ist die Angleichung der Amplituden der gerad- und ungeradzahligen Teiltöne jedoch noch nicht ganz ausgeprägt. Dagegen ist bei der Tonhöhe $h^1$ oberhalb des dritten Teiltons die gleichmäßige Abnahme aller Teiltöne klar zu erkennen (Abb. 2-7). Gleichzeitig ist die schwache Ausprägung des zweiten Teiltons mit deutlich über 10 dB Differenz zu den benachbarten Teiltönen noch deutlich vorhanden. Die leichte Welligkeit des Spektralverlaufs zu höheren Frequenzen ist eventuell durch Resonanzeigenschaften des Instruments zu begründen. So ist auch hier eine leichte Ausprägung des Formanten zwischen 3000 und 4000 Hz sichtbar.[39] Geräuschanteile spielen für alle Stimuli der Klarinette kaum eine Rolle. Die stärksten Anteile liegen etwa 60 dB unterhalb der stärksten Teiltonamplituden, womit ihr Einfluß als sehr gering eingeschätzt werden dürfte.

38 Meyer (1999). S. 66.

39 Laut Meyer befindet sich bei etwa $a^1$ die Überblasgrenze der B-Klarinette. Dieser Ton stellt für den Instrumentalisten einen schwerer zu spielenden Ton dar und drückt sich durch einen im Vergleich zu den umliegenden Tonhöhen schwächeren Klang aus. Vor der Stimuli-Aufnahme konnte mit der Klarinettistin geklärt werden, daß bei ihrem Instrument die Tonhöhe $h^1$ unbeeinflußt von o.g. Effekten war.

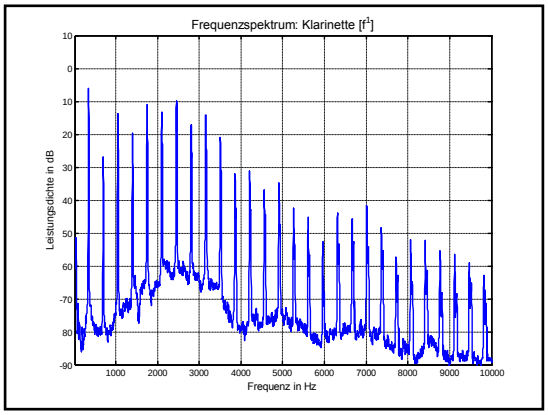


Abb. 2-6

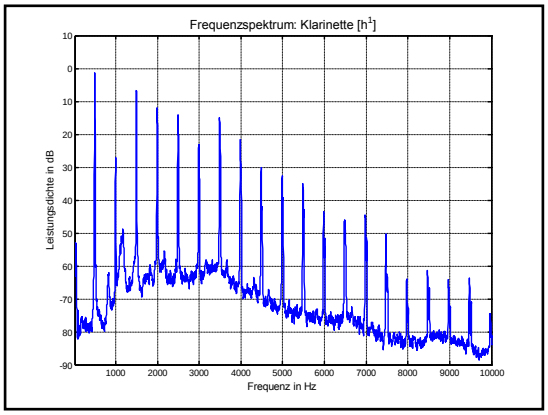


Abb. 2-7

Neben den Spektraleigenschaften der Stimuli sollen auch besondere Charakteristika bezüglich der Einschwingvorgänge der Instrumente betrachtet werden. Der bei den Stimuli-Aufnahmen einheitlich gewünschte weiche Ansatz hat laut Meyer bei der Klarinette die Auswirkung, daß die einzelnen Teiltöne zeitversetzt nacheinander einschwingen, d.h. von tieferen zu den höheren Frequenzbereichen. Die Gesamtdauer des weichen Ansatzes kann bei der Klarinette bis über 50 msec reichen. Wie in der nach Teiltönen getrennten Darstellung der Einschwingvorgänge in den Abb. 2-8 bis Abb. 2-11 zu sehen[40], trifft eine Einschwingdauer von circa 50 msec für alle Tonhöhen zu. Lediglich bei der Tonhöhe $f^1$ gibt es zwischen 50 und 175 msec eine weitere graduelle Zunahme der Amplitude. Diese Tatsache kann jedoch im Hinblick auf die Interpretation des MDS-Ergebnisses ignoriert werden, da $f^1$ nicht als Versuchs-Stimulus zum Einsatz kam.

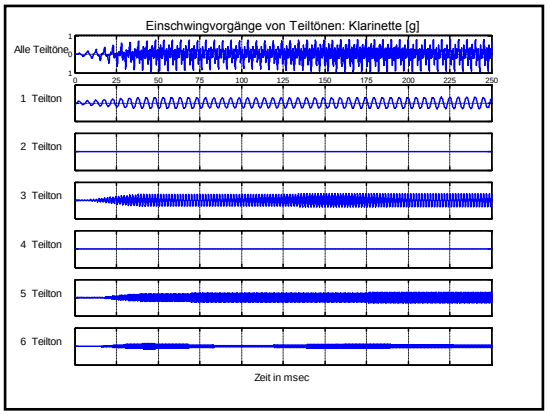


Abb. 2-8

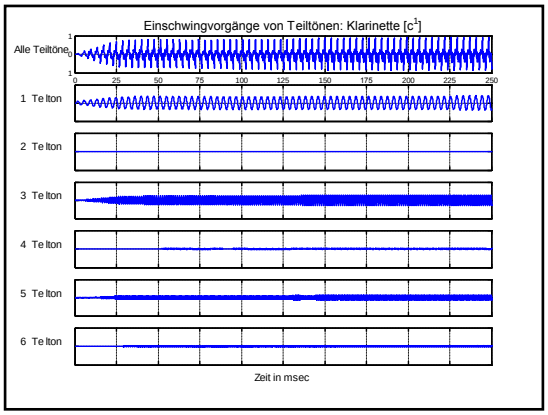


Abb. 2-9

40 Die zeitliche Darstellung einzelner Teiltöne wurde mittels digitaler Bandpassfilterung (Butterworth-Filter 3. Ordnung) umgesetzt. Der Durchlaßbereich beträgt einen Ganzton (relative Filterbreite). Die Mittenfrequenzen entsprechen den Teiltonfrequenzen. Sich durch das Filter ergebene Phasenverschiebungen der Zeitfunktionen führen zu zeitlichen Verschiebungen unterhalb 1 msec und können daher vernachlässigt werden. Die Berechnung und graphische Darstellung erfolgte mit *MATLAB*.

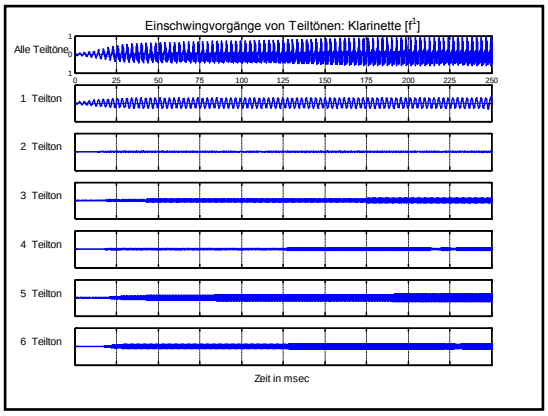


Abb. 2-10

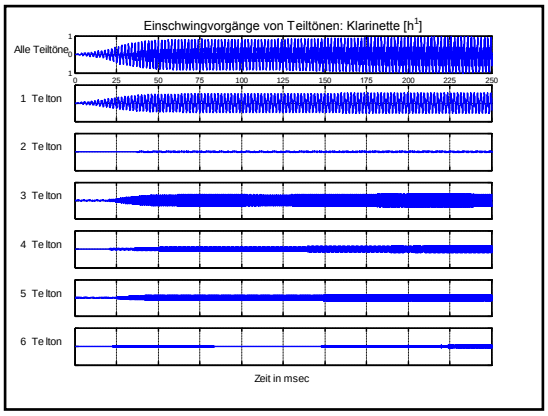


Abb. 2-11

Das sequentielle Einschwingen einzelner Teiltöne hin zu höheren Frequenzen ist bei den tieferen Tonhöhen stärker ausgeprägt. Während eine deutliche Zeitverzögerung zwischen Grundton und den weiteren Teiltönen von 10 bis 20 msec bei allen Tonhöhen zu erkennen ist, weisen die höheren Teiltöne untereinander kleinere Zeitverzögerungen auf und treten daher besonders bei den Tonhöhen $f^1$ und $h^1$ im Gesamteindruck nahezu synchron in Erscheinung. Nebenher läßt sich bei den tieferen Tonhöhen auch deutlich die kaum vorhandene Ausbildung des zweiten bzw. auch vierten Teiltons erkennen.

### 2.4.2 - Horn

Die Horn-Stimuli wurden über alle Tonhöhen in zwei Klangvarianten aufgenommen. Einmal in der gewöhnlichen Spieltechnik, bei der sich eine Hand teilweise im Trichter befindet; zweitens mit einem konischen Dämpfer aus Holz, welcher im Trichter befestigt wurde.

*a) Horn ohne Dämpfer*

Laut Meyer[41] besitzen Hörner zwei grundlegend unterschiedliche Spektralcharakteristiken. Während der tiefere Bereich deutlich vom Auftreten mehrerer Formanten geprägt ist, zeichnet die höheren Frequenzbereiche eine harmonische Obertonstruktur aus. Der Übergang zwischen beiden Gruppen erfolgt etwa bei der Tonhöhe $c^1$.

Beim Deutschen Doppelhorn gibt es laut Meyer neben einem Hauptformanten, weitere drei bzw. (bei Verwendung des F-Horns) vier Nebenformanten. Der Hauptformant liegt bei etwa 340 Hz und bildet unabhängig von der Lage der einzelnen Teiltöne im Spektrum das Maximum. Wie für die Stimuli der Tonhöhen g und $c^1$ in Abb. 2-12 und 2-13 gesehen werden kann, ist keine Ausprägung des Hauptformanten zu erkennen. Offensichtlich liegen diese Tonhöhen bzw. ihr Grundton mit etwa 197 Hz bereits zu hoch, um zu einer Ausprägung dieses Formanten zu führen, da die Teiltonabstände bereits keine Abbildung des Formanten mehr ermöglichen. Demnach verlagert sich jedoch das Maximum auf den ersten Nebenformanten, welcher getreu Meyers Angaben bei etwa 725 Hz zu finden ist. Da bei der Tonhöhe g der zweite und dritte Teilton in den Bereich dieses Formanten fallen, bilden diese auch das Maximum des Teiltonspektrums. Auch bei der Tonhöhe $c^1$ wird der Grundton bei ca. 263 Hz noch vom im Formantbereich liegenden zweiten Teilton überragt. Der

41 Meyer (1999). S. 45-49.

Grundton liegt für beide Töne jedoch nur etwa 5 dB unterhalb des Maximums. Für beide Tonhöhen sind weitere, weniger ausgeprägte Formanten zwischen 2000 und 3000 Hz, sowie bei etwa 4000 Hz zu verzeichnen. Diese Formantlagen weichen von denen von Meyer angegebenen ab, welche mit 1225, 2000, sowie beim F-Horn zudem 3500 Hz angegeben werden. Diese Abweichung läßt sich möglicherweise damit begründen, daß sich beide Tonhöhen im oberen Bereich des formantgeprägten Bereichs befinden, wo der absolute Freqenzabstand zwischen den Teiltönen bereits keine genaue Abbildung der Feinstruktur des Formantverlaufs zuläßt und somit eventuelle, real existierende Maxima bei den von Meyer genannten Frequenzen nicht abgebildet werden können. Außerdem können zusätzlich instrumentenspezifische Resonanzeigenschaften diese Differenz verursachen. Mit Zunahme der Lautstärke verschieben sich laut Meyer die Formantbereiche geringfügig zu höheren Frequenzen. Gegen die Annahme dieser Feststellung als Begründung für die Differenz spricht jedoch das Vorliegen einer eher geringen Dynamikstufe bei der Aufzeichnung.

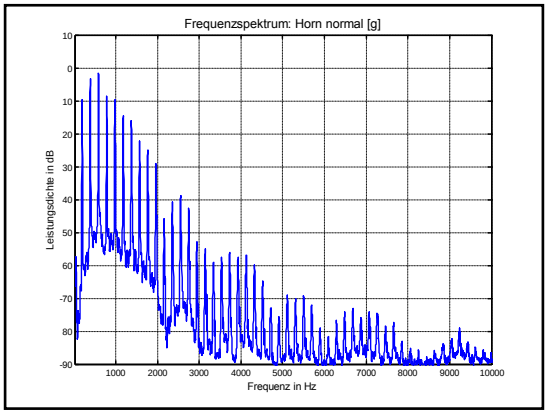


Abb. 2-12

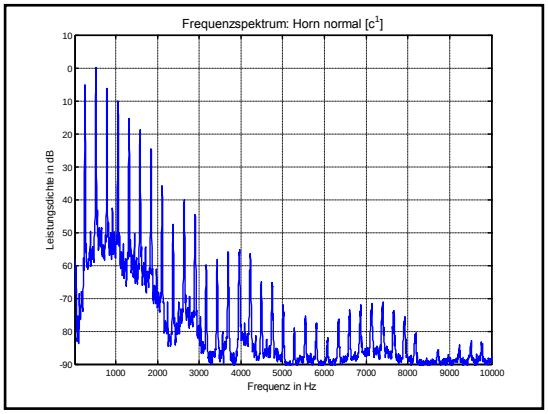


Abb. 2-13

Die Tonhöhe $c^1$ stellt, wie bereits erwähnt, den ungefähren oberen Grenzbereich des Formanteinflusses dar, wobei sich beim vorliegenden Stimulus noch ein deutlicher Einfluß der Formanten abzeichnet. Dagegen ist dieser Einfluß bereits bei der Tonhöhe $f^1$ nur noch in dem charakteristischen Spektralverlauf des Formanten bei 725 Hz zu erkennen. Auch tritt hier bereits der bei ca. 350 Hz liegende Grundton (Abb. 2-14) als stärkster Teilton in Erscheinung, während weitere Nebenformanten bereits nicht mehr zu erkennen sind. Das Spektrum der Tonhöhe $h^1$ (Abb. 2-15) ist geprägt von einer nahezu idealen Obertonstruktur, ohne nennenswerten Formanteinfluß.

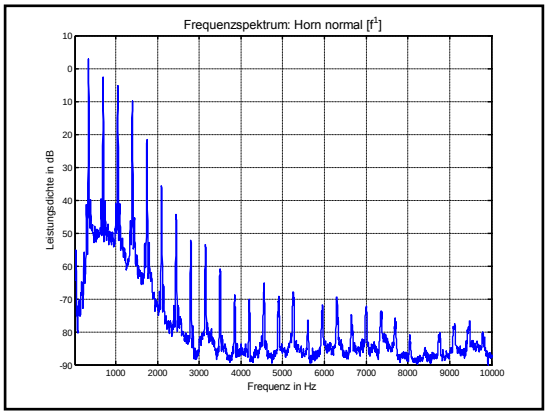


Abb. 2-14

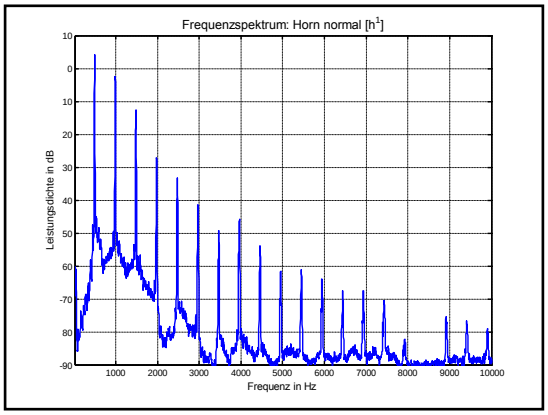


Abb. 2-15

Der Einfluß von Geräuschanteilen ist Meyer nach bei Hörnern aufgrund ihrer schwachen Ausprägung zu vernachlässigen. Laut Meyer reichen Geräuschanteile bis etwa 3000 Hz, wie es auch aus den Spektren aller Stimuli nachzuvollziehen ist. Sie liegen circa 55 dB unterhalb der stärksten Teiltöne und ihr Vorkommen ist womöglich damit zu begründen, daß Hornklänge in einer leiseren Dynamikstufe aufgezeichnet wurden, womit eine geringerer Abstand zwischen tonalen und geräuschhaften Anteilen erreicht wird als bei lauteren Dynamikstufen.

Beim Horn treten vor allem bei härterem Blasansatz sogenannte Vorläuferimpulse auf. Beim für die Stimuli gewünschten weicheren Ansatz ist der Einfluß des Vorläuferimpuls weniger bedeutend bzw. tritt er laut Meyer klanglich in den Hintergrund. Die Einschwingdauer kann bei angestoßenen Tönen oberhalb der Tonhöhe f 30 bis 40 msec betragen, bei weichem Ansatz jedoch über 100 msec andauern. Über eventuelle Differenzen zwischen den zeitlichen Einschwingvorgängen der einzelnen Teiltöne bzw. Frequenzregionen macht Meyer keine Angaben. Bei den Stimuli ist ein zeitlich sehr ausgedehnter Einschwingvorgang für alle Tonhöhen zu beobachten (Abb. 2-16 bis 2-19). Erst bei etwa 250 msec ist die Amplitudenstärke des quasi-stationären[42] Bereichs erreicht.

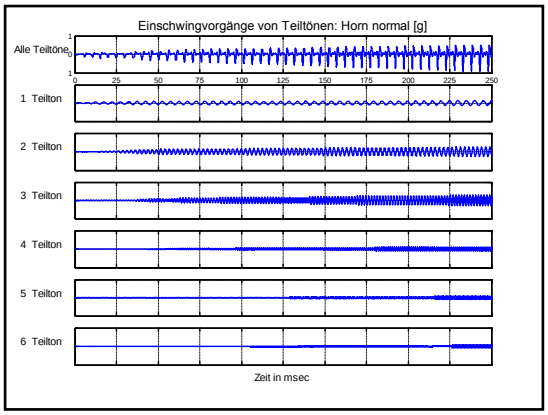


Abb. 2-16

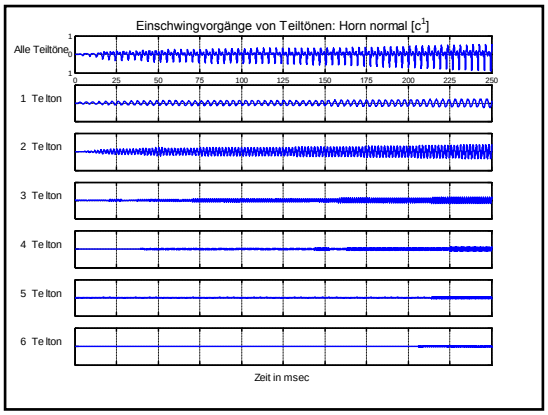


Abb. 2-17

42 Aufgrund zeitlicher Fluktuation in beispielsweise der Lautstärke und dem Spektrum von akustischen Instrumentalklängen ist die im Idealfall *stationäre* Phase nach dem Einschwingvorgang in Realität als *quasi-stationär* einzustufen.

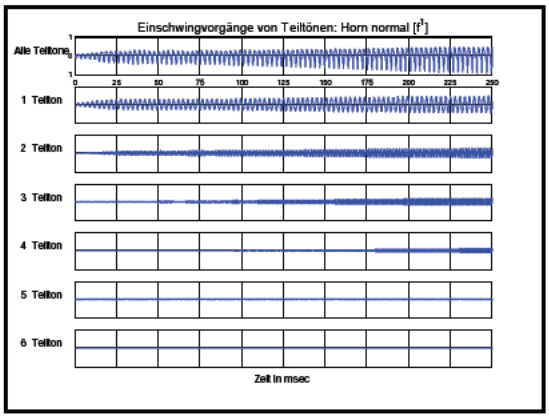


Abb. 2-18

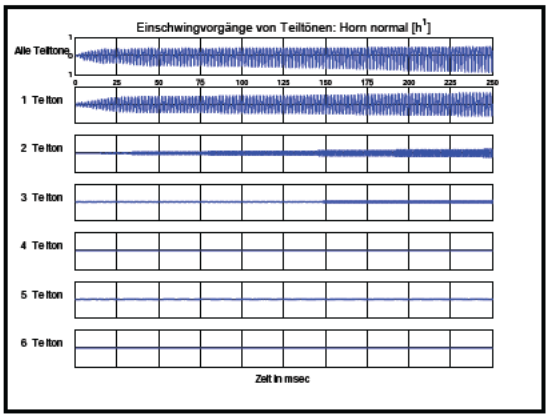


Abb. 2-19

Bezüglich der zeitlichen Relationen einzelner Teiltöne zueinander ist das sequentielle Einschwingen des nächsthöheren Teiltons sehr ausgeprägt. Der Zeitversatz zum nächsten einsetzenden Teilton erstreckt sich in einem Bereich von 10 bis 100 msec. Weiterhin ist eine Tendenz der Zunahme des Zeitversatzes für höhere Teiltöne zu erkennen. Jedoch ist eine Schlußfolgerung auf systematische Gesetzmäßigkeiten bei der geringen Anzahl von Beispielen nicht ratsam, da diese Differenzen ebensogut aus Variabilitäten im Blasansatz oder der Dynamik herrühren können. Die Existenz von breitbandigen Vorläuferimpulsen deutet sich nur bei der tiefsten Tonhöhe g an. Wie in dem Sonagram in Abb. 2-20 für die ersten 150 msec zu sehen ist, gibt es zu Beginn zwischen 250 und 750 Hz neben den eigentlichen Teiltönen zusätzliche Frequenzanteile. Gegenüber den Teiltönen sind sie offensichtlich schwächer ausgeprägt, womit sie als eher schwache bis mittelstarke Vorläuferimpulse aufgefaßt werden können. Somit kann ein zu berücksichtigender Einfluß durch Vorläuferimpulse an dieser Stelle auf die Tonhöhe g eingegrenzt bzw. für die restlichen Tonhöhen ausgeschlossen werden.

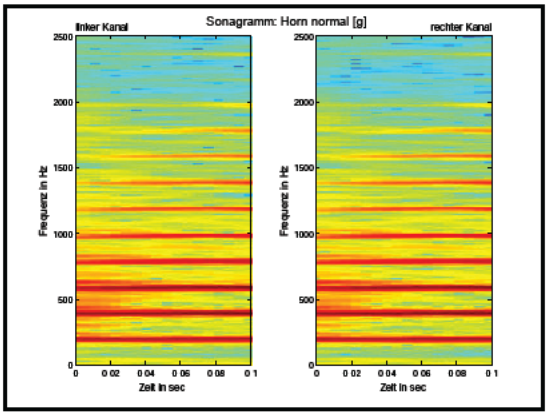


Abb. 2-20

*b) Horn mit Dämpfer*

Zum Einfluß eines Dämpfers für das Horn gibt es bei Meyer keine Auskunft. Als einzige spezielle Spielweisen sind das nahezu vollständige Verschließen des Trichters mit der Hand, das sogenannte *Stopfen*, und das Spielen mit vollständig offenem Trichter genannt. Während erstere Technik Ähnlichkeiten mit der Wirkung eines Dämpfers aufweist, finden sich jedoch die von Meyer hierzu genannten Charakteristiken, bis auf die Ausweitung des Frequenzbereichs durch ausgeprägtere hohe

Teiltöne, nicht in den Spektren der Stimuli dieser Arbeit wieder. Allerdings wird bei der Besprechung der Trompete auf verschiedene Dämpfertypen eingegangen.[43] Da für Trompeten gewöhnlich keine Dämpfer aus Holz verwendet werden, läßt sich eine Gemeinsamkeit am ehesten auf die geometrische Form zurückführen, die beim verwendeten hölzernen Dämpfer des Horns konusförmig ist. Für den normalen, konischen Dämpfer für die Trompete nennt Meyer als maßgebliche, den Klang variierende Charakteristiken die deutliche Dämpfung des Grundtons im Spektrum und eine Zunahme der Amplitudenstärke oberhalb 4000 Hz.

Wie in Abb. 2-21 zu sehen ist, kann diese schwächere Ausprägung beim tiefsten Ton g zwar erkannt werden, nur resultiert sie letztlich wahrscheinlich nicht aus dem Einfluß des Dämpfers allein. Hierfür spricht, daß bei den übrigen Tonhöhen der Grundton immer mit am stärksten ausgeprägt ist. Zudem gibt es deutliche Parallelen zu der Spektralcharakteristik der Tonhöhe g des Horns ohne Dämpfer. Daher kann davon ausgegangen werden, daß auch hier ein starker Einfluß von Formanten für die tieferen Tonlagen entscheidend ist. Beim genaueren Hinsehen überrascht es, Formantbereiche vorzufinden, die genau in die von Meyer genannten Frequenzregionen fallen, was im Vergleich beim Horn ohne Dämpfer nicht zutrifft. Neben dem stärksten Nebenformanten, welcher wie im ungedämpften Fall bei etwa 725 Hz angesiedelt ist, liegen die anderen etwa bei 1100, 2200 und 3500 Hz. Der wesentliche Unterschied zur Tonhöhe g für das Horn ohne Dämpfer liegt jedoch in dem deutlich obertonreicheren Spektrum. Während beim Horn ohne Dämpfer eine obere Grenze von 4000 Hz erreicht wird, erweitert sich der Frequenzbereich mit Dämpfer bis zu 7000 Hz.[44] Auch weist dieser erweiterte Frequenzraum weitere mehr oder weniger ausgeprägte Nebenformanten auf. Der Formanteinfluß prägt ebenso das Spektrum der Tonhöhe $c^1$. Jedoch reicht das Spektrum (Abb. 2-22) für diese Tonhöhe nur bis etwa 5500 Hz.

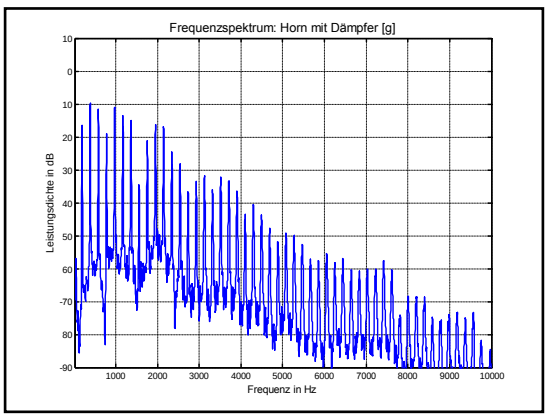


Abb. 2-21

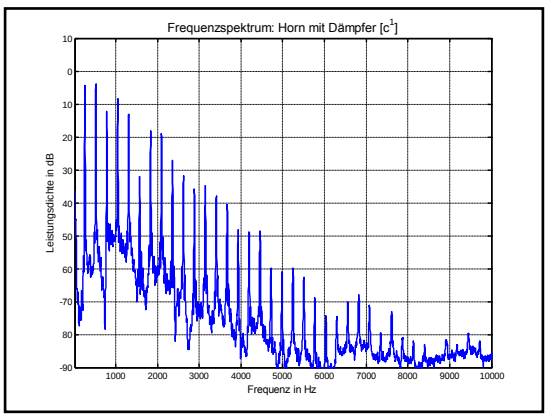


Abb. 2-22

43 Meyer (1999). S. 53.

44 Nicht alles in den Spektren abgebildete kann wegen verschiedener Einflüsse auch wahrgenommen werden. Für die Hörbarkeit der zunehmend leiseren hohen Frequenzen spielt der erzielte Dynamikbereich in der Abhörsituation eine Rolle. Für die Bestimmung von Obergrenzen für den noch wahrnehmbaren hochfrequenten Bereich ist es daher sinnvoll, sich auf einen realistisch annehmbaren Pegelschwellwert zu beziehen. Aufgrund der ruhigen Abhörumgebung und der mittleren Wiedergabelautstärke wird hierfür ein relativer Wert von 60 dB unterhalb des Spektralmaximums gewählt.

Oberhalb der von Meyer angegebenen Grenze, von wo an der Formanteinfluß geringer wird, läßt sich für die Tonhöhe $f^1$ wie beim ungedämpften Pendant noch ein leichter Einfluß der Formanten erkennen (Abb. 2-23). Durch das breitere Frequenzspektrum zeichnen sich jedoch insgesamt mehr Formanten ab als für den ungedämpften Fall. Für die Tonhöhe $h^1$ läßt sich bereits kaum noch ein Formanteinfluß erkennen und es bildet sich wie für den ungedämpften Fall der Verlauf einer harmonischen Obertonreihe aus (Abb. 2-24).

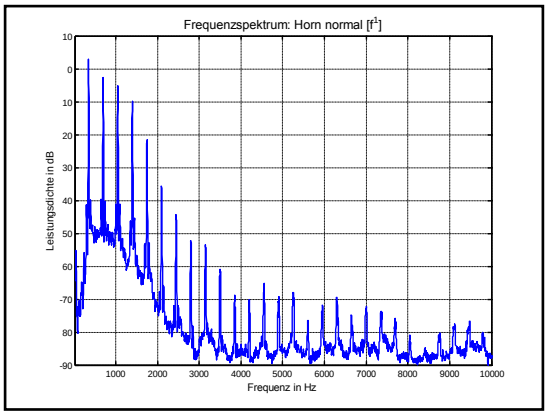


Abb. 2-23

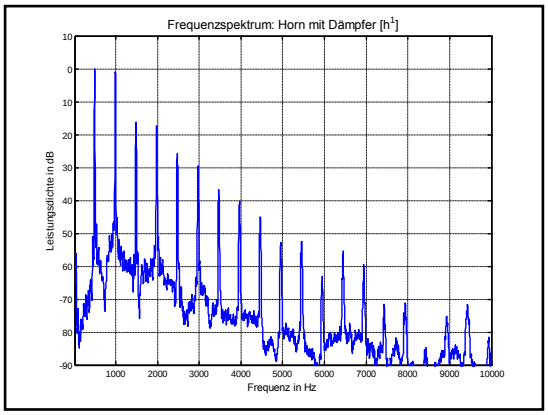


Abb. 2-24

Die Eigenschaft des konischen Dämpfers bei der Trompete, die eine Schwächung des Grundtons hervorruft, läßt sich wie gezeigt nicht auf den konischen Holzdämpfer übertragen; sehr wohl jedoch die zweite Eigenschaft, die Ausweitung des Frequenzbereichs hin zu höheren Frequenzen. Mit dieser Frequenzraumerweiterung geht der zuvor schon als gering eingestufte Geräuschanteil einher, welcher sich auch nach oberen Frequenzen hin ausweitet.

Für den Einschwingvorgang spielen durch den Dämpfer bedingte Einflüsse offensichtlich weniger eine Rolle. So ist die Einschwingdauer des Gesamtsignals mit etwa 250 msec etwa vergleichbar mit der ohne Dämpfer (Abb. 2-25 bis 2-28). Auch gibt es keine grundlegenden Unterschiede zwischen den zeitlichen Beziehungen der Teiltöne untereinander. Vorhandene Unterschiede resultieren womöglich nur aus Variabilitäten des Anblasens. Allerdings bleibt anzumerken, daß sich das reichere Spektrum des Horns mit Dämpfer nicht direkt von Beginn an einstellt, sondern erst bei Vollendung des Einschwingvorgangs die maximale Amplitude erreicht wird, da der Einsatz der höheren Teiltöne zum Grundton zeitversetzt ist. Wie beim Horn ohne Dämpfer läßt sich auch hier ein Vorläuferimpuls lediglich bei der Tonhöhe g erkennen. Im Vergleich zum Horn ohne Dämpfer ist dieser Vorläuferimpuls breitbandiger und dadurch stärker ausgeprägt (Abb. 2-29). Dies wird im MDS-Ergebnis zu berücksichtigen sein.

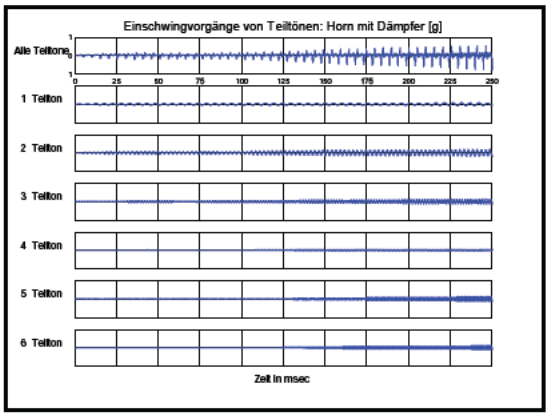


Abb. 2-25

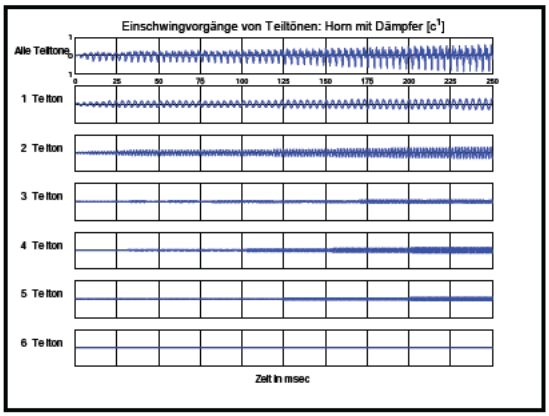


Abb. 2-26

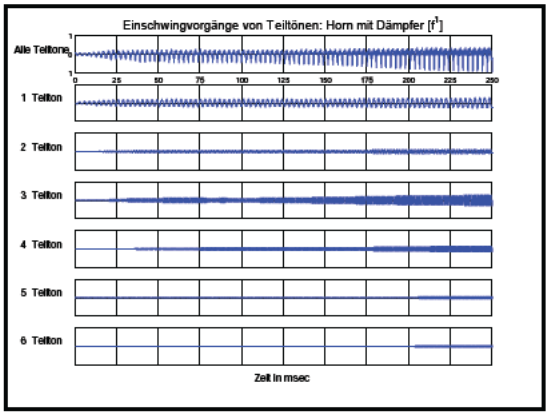


Abb. 2-27

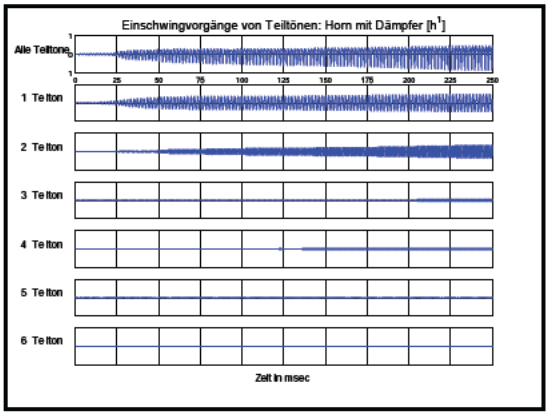


Abb. 2-28

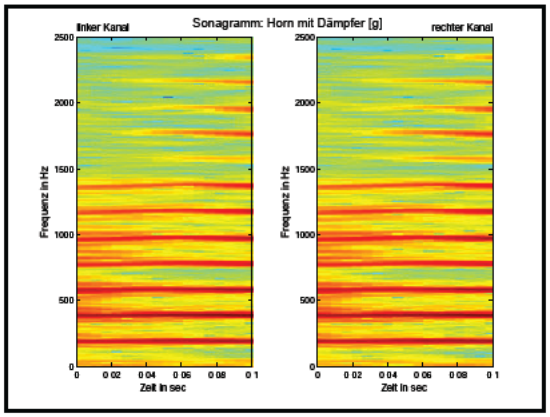


Abb. 2-29

### 2.4.3 - Baßflöte

Für die Baßflöte liegen Meyer keine Ergebnisse über die Spektral- und Einschwingcharakteristik vor, da sich seine Betrachtungen auf die große C-Flöte und Piccolo beschränken.[45] Da die Klangerzeugung und akustische Wirkungsweise der tieferen Vertreter der Querflötenfamilie jedoch untereinander vergleichbar sind, werden infolge für die große Querflöte geltende Charakteristiken mit Zuhilfenahme der an den Stimuli vorgenommenen physikalisch-akustischen Analysen zu Rate gezogen. Querflöten

45 Meyer (1999). S. 59-62.

zeichnet in der Regel eine sehr gleichmäßige Obertonstruktur aus, welche bis auf die tiefsten Töne über alle Frequenzbereiche zu erwarten ist. Meyer nach läßt sich jedoch durch zahlreiche Möglichkeiten der Anblastechnik die Ausprägung der untersten Teiltöne „in verhältnismäßig weiten Grenzen variieren“[46]. Zudem kommt es durch die für Querflöten spezifische Anblastechnik im Vergleich zu anderen Blasinstrumenten zu einem relativ hohen Anteil an Geräuschkomponenten, welche den typischen Querflötenklang ausmachen. Dieses kontinuierliche Geräusch entsteht aus dem durch das Blasen verursachten Rauschen. Gleichzeitig können durch dieses breitbandige Rauschen zusätzlich Korpusresonanzen angeregt werden, welche bei überblasenen Tönen[47] auch zur Anregung von Sub-Harmonischen führen können, d.h. Oktavtönen unterhalb des ersten Teiltons und seiner Vielfachen.

In den dargestellten Spektren der aufgenommenen Baßflöten-Stimuli weisen die Tonhöhen g und $c^1$ (Abb. 2-30 und 2-31) vergleichbare Charakteristiken auf. Der Grundton (etwa bei 197 Hz bzw. 263 Hz) ist in beiden Fällen deutlich als stärkster Teilton ausgeprägt. Es lassen sich fünf bis sieben weitere Teiltöne erkennen, welche jedoch nicht den Verlauf einer idealen Obertonreihe besitzen, da keine streng-stetige Abnahme der Amplitude zum nächsthöheren Teilton zu beobachten ist. Oberhalb der ausgeprägten Teiltöne verbleiben lediglich Geräuschanteile im Spektrum, mit einem Abstand von etwa 50 dB zu den stärksten Teiltönen.

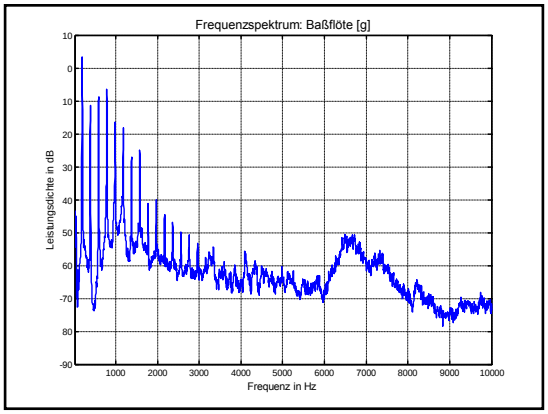


Abb. 2-30

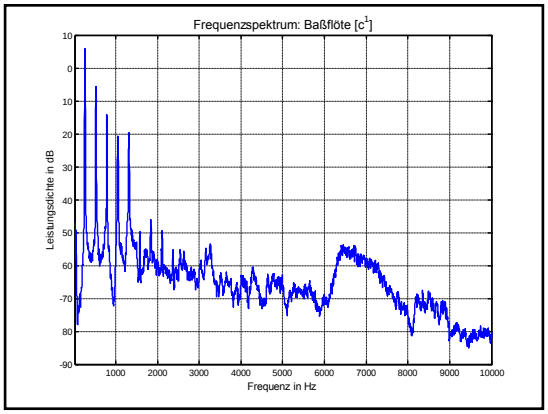


Abb. 2-31

Die Spektren der Tonhöhen $f^1$ und $h^1$ sind in Abb. 2-32 und 2-33 abgebildet. Die Gestalt der Obertonreihe ist weitestgehend mit denen der beiden tieferen Tonhöhen vergleichbar. Auch hier gibt es nur eine geringe Anzahl deutlich ausgeprägter Teiltöne. In beiden Fällen werden hier jedoch zusätzlich auch Sub-Harmonische des Grundtons (bei etwa 175 Hz bzw. 249 Hz), sowie die ersten beiden ungeradzahligen Vielfachen (mit 3- und 5-facher Frequenz) dieser Sub-Harmonischen angeregt. Die Amplituden dieser zusätzlich angeregten Resonanzen liegen etwa 30 dB unterhalb der stärksten Teiltöne. Auch für $f^1$ und $h^1$ verbleibt oberhalb der Teiltöne nur der Geräuschanteil im Spektrum. Der Geräuschanteil ist für alle Tönhöhen spektral sehr ähnlich verteilt und weist zwischen 6000 und 7000 Hz ein deutliches Maximum auf, was als ein Formant aufgefaßt werden könnte. Erwähnenswert bzw. für die spätere Interpretation zu berücksichtigen ist die Tatsache,

46 Meyer (1999). S. 59.
47 Beim Überblasen von Tönen werden höhere Oktaven (zur Grundstimmung) angeregt.

daß $c^1$ gerade im Vergleich zu den anderen Versuchs-Stimuli g und $h^1$ einen deutlich geringeren Frequenzbereich der Teiltöne aufweist und sich nur bis etwa 1500 Hz eine stetige Obertonstruktur erkennen läßt. Dies ist möglicherweise durch eine bei der Aufnahme geringere Dynamik oder einer leicht abweichenden Anblastechnik bedingt. Nach Meyer üben bei Querflöten Unterschiede in der Dynamikstufe auf die spektrale Frequenzausdehnung einen sehr starken Einfluß aus.

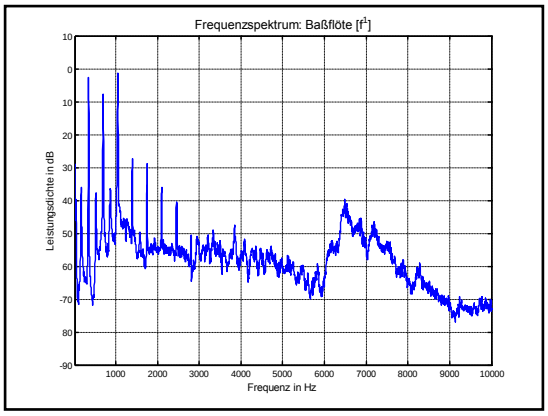


Abb. 2-32

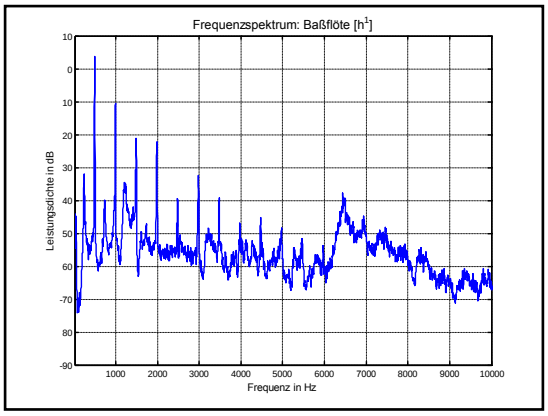


Abb. 2-33

Die Querflöte zeichnet sich im Einschwingvorgang dadurch aus, daß sie in der Regel von allen Blasinstrumenten die längste Einschwingdauer besitzt. Je nach Art des Blasansatzes ist zudem das Vorhandensein bzw. die Rolle eines Vorläufertons zu berücksichtigen. Bei einem wie hier vorliegenden eher weichen Ansatz ist zu erwarten, daß Vorläufertöne, sofern überhaupt vorhanden, weniger stark ausgeprägt sind. Nennenswert ist zudem die Abhängigkeit des Einschwingverhaltens einer Querflöte von der Tonlage. So existieren in den verschiedenen Tonlagen unterschiedliche zeitliche Relationen zwischen den einzelnen Teiltönen, welche letztlich auch die Vorläufertöne verursachen. Die von Meyer für die einzelnen Tonlagen genannten Charakteristiken können jedoch nicht ohne weiteres auf die Baßflöte übertragen werden. Daher bedarf es einer genaueren Untersuchung der temporalen Eigenschaften der entsprechenden Stimuli.

Betrachtet man die Einschwingvorgänge der Stimuli (Abb. 2-34 bis 2-37), gibt es für alle Tonhöhen eine einheitliche Einschwingdauer von etwa 150 msec mit einer weiteren leichten Amplitudenzunahme über die folgenden 100 msec. Somit ist, verglichen mit allen Stimuli, nicht die Flöte das Instrument mit der längsten Einschwingdauer, sondern in diesem Fall das Horn. Für jede Tonhöhe schwingt des weiteren der Grundton vor allen anderen Teiltönen ein. Dies widerspricht zumindest der von Meyer für die C-Flöte genannten Eigenschaft, daß im tiefen Tonbereich der zweite Teilton (Oktavton) zuerst einschwingt. Als Vergewisserung, daß keine breitbandigen Frequenzanteile beim Einschwingvorgang vorhanden sind, wurden wie beim Horn Sonagramme analysiert. Aus ihnen lassen sich keine Vorläufertöne oder andere breitbandigen Frequenzanteile erkennen.[48]

48 Daher wird auch hier auf die Sonagramm-Darstellung verzichtet.

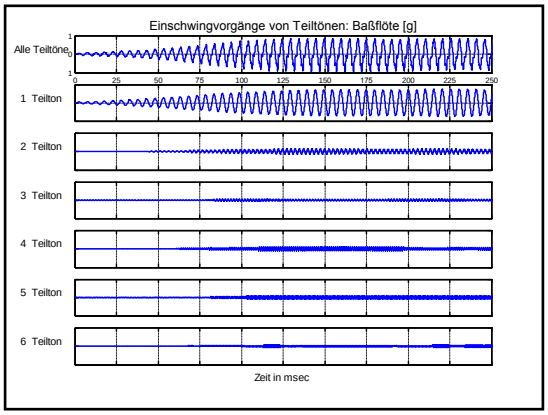


Abb. 2-34

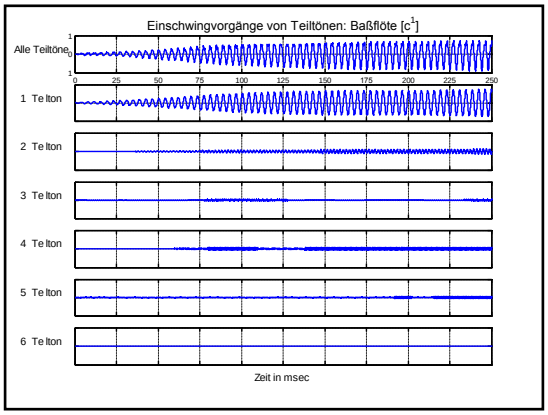


Abb. 2-35

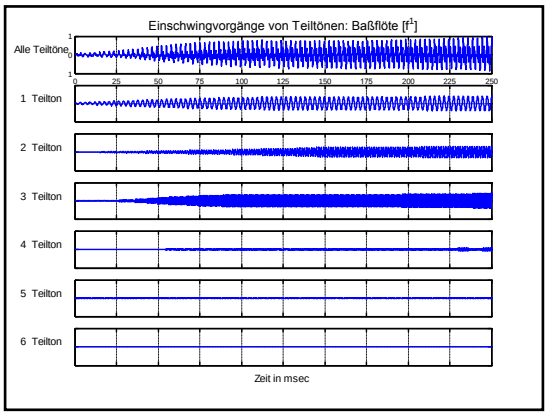


Abb. 2-36

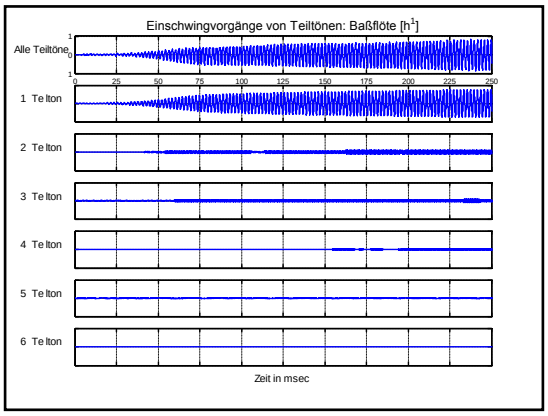


Abb. 2-37

Auch läßt sich über die verschiedenen Tonhöhen keine Änderung in den zeitlichen Relationen der Teiltöne untereinander (wie bei der großen C-Flöte) erkennen. Somit weist die Baßflöte zumindest im Tonbereich der Stimuli keine typischen Eigenschaften des höheren Vertreters der Querflötenfamilie auf. Bezüglich dieser zeitlichen Beziehungen setzt für die tiefste Tonhöhe g der zweite Teilton etwa 50 msec nach dem Grundton ein. Es folgen zeitversetzt jedoch, in erheblich geringeren Abständen, die restlichen höher gelegenen Teiltöne, welche nahezu gleichzeitig einsetzen. Bei den Tonhöhen $c^1$ und $h^1$ läßt sich eine tendenzielle Entwicklung erkennen, in welcher zunehmend weniger Obertöne vorhanden sind. Der höchste Oberton folgt hierbei jeweils um 100 bis 150 msec (respektive $c^1$ und $h^1$) deutlich abgetrennt von den anderen, dagegen nahezu synchron auftretenden, Teiltönen. Auch hieraus darf aufgrund der geringen Anzahl von Beispielen keine generelle Gesetzmäßigkeit abgeleitet werden. Die Tonhöhe $f^1$ weicht in puncto der oben genannten Eigenschaften von den anderen Stimuli ein wenig ab. Sie spiegelt keine der bereits genannten Regelmäßigkeiten zwischen temporalen und spektralen Eigenschaften ideal wider, da sie beispielsweise eine relativ starke Ausprägung des dritten Teiltons besitzt. Diese Abweichung resultiert, wie sich auch durch subjektiven Hörvergleich feststellen läßt, mit einiger Sicherheit aus einer von den anderen Flötenklängen etwas abweichenden Anblastechnik. Doch kann auch hier diese Tatsache für die Interpretation des MDS-Ergebnisses ignoriert werden, da die Tonhöhe $f^1$ keinen Versuchs-Stimulus darstellt.

## 2.5 - Vergleich der Instrumentalklänge

Zur weiteren Betrachtung und späteren Interpretation der MDS ist es sinnvoll, zusammenfassend grundlegende Unterschiede bzw. Ähnlichkeiten zwischen den in den Stimuli verwendeten Instrumenten hervorzuheben. Absolute Ausprägungen bestimmter Charakteristiken sind hierbei weniger entscheidend, als die relativen Unterschiede zwischen Stimuli bzw. verschiedenen Instrumenten, da es sich bei den paarweisen Ähnlichkeitsbewertungen gleichermaßen um Relativbewertungen handelt und auch die geometrische Konstellation des MDS-Ergebnisses relative Unterschiede zwischen Stimuli abbildet. Dieser zusammenfassende Vergleich läßt sich analog zur Betrachtung der einzelnen Instrumente nach Unterschieden in den spektralen und temporalen Eigenschaften aufgliedern.

### 2.5.1 - Unterschiede im Spektrum

Der deutlichste auf das Spektrum zurückzuführende Unterschied zwischen den Instrumenten bzw. Instrumentalvariationen ist die frequenzmäßige Ausdehnung der Teiltonstruktur. Diese Charakteristik vermag es, Unterschiede zwischen allen vier Instrumenten bzw. Instrumentalvariationen zu verdeutlichen. Selbst bei den auf dem selben Instrument beruhenden Varianten des Horns, resultiert die Nutzung eines Dämpfers in der deutlichen Ausweitung der Bandweite der Obertonstruktur. Zwischen den Instrumenten existieren auch klare Unterschiede hinsichtlich des Reichtums ihrer Obertonstruktur. Während die Klarinette und die beiden Horn-Varianten eine bis in hohe Frequenzen deutlich ausgeprägte Obertonstruktur aufweisen, besitzt die Baßflöte dagegen ein auf lediglich sechs bis sieben Teiltöne beschränktes Spektrum. Darüber hinaus besitzt die Klarinette gerade für die tieferen Tonhöhen durch das Fehlen geradzahliger Teiltöne eine prägnante, einzigartige Spektralcharakteristik.

Das Vorhandensein von Formanten und Unterschiede in ihrer Lage führen auch zu Unterschieden in der Spektralcharakteristik der Instrumente. Vor allem die Horn-Stimuli sind von der Existenz von Formanten deutlich geprägt: das Horn mit Dämpfer wegen seines hochfrequenteren Spektrums sogar mehr als das ungedämpfte Horn. Interessant ist auch die durch den Dämpfer bedingte, verschobene Lage der Formanten. Allerdings läßt beim Horn der Einfluß der Formanten mit ansteigender Tonhöhe nach, so daß einerseits die höchsten Töne den tieferen in dieser Beziehung unähnlicher werden, andererseits jedoch auch den Tönen der anderen Instrumente (mit geringem Formantvorkommen) wiederum ähnlicher werden könnten. Die Klarinette weist zwar Anzeichen für Formantausprägung auf, jedoch sind diese über alle Tonhöhen nicht gleichstark ausgeprägt und keinesfalls über alle Tonhöhen gleich positioniert, so daß ihnen letztlich keine große Bedeutung zukommen dürfte. Prägend für die Klarinette ist primär ihre schon erwähnte spezielle Spektralcharakteristik, die sich anders als bei Formanten (mit ihren absoluten Frequenzbereichen) relativ zur Grundtonfrequenz verhält und somit ihre Position frequenzabhängig ist. Die Baßflöte besitzt nur einen Formantbereich. Dieser ausgeprägte Formant (bei 6000-7000 Hz) ist jedoch bei allen Tonhöhen zu verzeichnen. Angesichts der relativ begrenzten Obertonstruktur könnte dieser oberhalb der Teiltöne angesiedelte Formant durchaus eine charakteristische Stellung einnehmen.

Ein letzter zu berücksichtigender Aspekt der Spektralcharakteristik sind das Vorkommen und die Ausprägung von Geräuschanteilen im Klang. Besonders die Querflötenfamilie ist für ihren

charakteristischen Geräuschanteil bekannt. Bei diesen Stimuli ist ein minimaler Abstand des Rauschens zu den stärksten Teiltonanteilen (‚Signal-Rausch-Abstand‘) von etwa 50 dB zu verzeichnen. Noch wichtiger ist jedoch, daß sich der Geräuschanteil über die Frequenz bis etwa 8000 Hz ausdehnt, d.h. einschließlich des Formanten. Im Vergleich besitzen beide Horn-Varianten einen minimalen Abstand zum Geräuschanteil von etwa 55 dB. Zusammen mit einer geringeren Frequenzausdehnung bis nur 3000 bzw. mit Dämpfer bis 4000 Hz, führt dies zu einem deutlich weniger vom Geräusch geprägten Klang des Horns. Bei den Spektren der Klarinetten-Stimuli ist zwar ein Geräuschanteil bis etwa 4000 Hz zu erkennen, doch liegen die stärksten Anteile 60 dB unterhalb der stärksten Teiltöne. Zusammen mit der starken und breitbandigen Ausprägung der Obertonstruktur ist für die Klarinette der geringste Einfluß von Geräuschanteilen anzunehmen. Für Horn und Klarinette ist ferner anzunehmen, daß aufgrund des begrenzten Dynamikbereichs der Abhörsituation, die Geräuschanteile bereits an der Schwelle der Wahrnehmbarkeit liegen.

### 2.5.2 - Unterschiede im Einschwingverhalten

Für alle Instrumente wurde bei der Aufnahme ein weicher Blasansatz verwendet um eine Vergleichbarkeit der Transienten ermöglichen bzw. ihre Variabilität einschränken zu können. Aus instrumententechnischen Gründen führt ein weicher Blasansatz nicht zwangsläufig zu einem identischen Einschwingvorgang. Die wohl offensichtlichste Charakteristik eines Einschwingvorgangs ist ihre Dauer, vom erstmaligen Anschwingen bis zum Erreichen des quasi-stationären Zustands. Hier treten zwischen den Instrumenten auch die deutlichsten temporalen Unterschiede auf. Mit einer Einschwingdauer von 250 msec erreichen die Horn-Stimuli den quasi-stationären Zustand deutlich als letzte. Ein Unterschied in der Einschwingdauer zwischen den beiden Instrumentalvariationen des Horns läßt sich nicht erkennen. Die Baßflöte folgt mit einer Einschwingdauer von etwa 150 msec, zusätzlich eines weiteren, geringfügigen Ansteigens der Amplitude in den darauffolgenden 100 msec. Somit sind für das Horn und die Baßflöte die Einschwingdauern etwa vergleichbar. Einen sehr deutlichen Unterschied besitzen jedoch Horn und Baßflöte zur Klarinette, die mit einer Dauer von circa 50 msec mindestens dreimal schneller einschwingt. Dieser Unterschied wird bei der Interpretation des MDS-Ergebnisses auf jeden Fall zu berücksichtigen sein. Allgemein läßt sich außerdem anmerken, daß die Einschwingdauern weitestgehend unabhängig von der Tonhöhe sind.

Bezüglich der zeitlichen Einschwingvorgänge einzelner Teiltöne läßt sich für alle Instrumente ein mehr oder weniger ausgeprägtes sequentielles Einschwingen der Teiltöne (von tieferen zu höheren) feststellen. Diese Tatsache ist nicht selbstverständlich, da wie bereits besprochen bei der großen C-Flöte in manchen Tonbereichen höhere Teiltöne zuerst einschwingen. Diese Eigenschaft bestätigt sich hier jedoch nicht und kann daher außer Acht gelassen werden. Wie festgestellt, variiert das sequentielle Einschwingen von Instrument zu Instrument, doch kann dieser Einfluß als wenig bedeutsam eingestuft werden. Denn Unterschiede zwischen den Teiltönen, oberhalb des ersten, lassen wegen relativ kleiner Unterschiede, keiner konsistenten Tendenz und der geringen Anzahl von Beispielen nur schwer Rückschlüsse auf systematische Gesetzmäßigkeiten zu. Nennenswerte Unterschiede und Ähnlichkeiten sind diesbezüglich jedoch die gemeinsame Eigenschaft der Klarinette und der

Baßflöte, daß der erste Teilton zeitlich von den anderen deutlich abgetrennt ist. Bei der Klarinette sind es 10 bis 20 msec Versatz, bei der Baßflöte aufgrund der längeren Einschwingdauer entsprechend 50 msec. Die Gruppe von Teiltönen oberhalb des ersten schwingt in einigen Fällen annähernd synchron ein. Gegenüber der Klarinette und Baßflöte besitzt das Horn eine recht ausgewogene zeitliche Verteilung des Eintretens der Teiltöne. Zudem ist ihr sequentielles Einschwingen zeitlich sehr ausgedehnt, wobei einzelne Verzögerungen zwischen Teiltönen bis zu 100 msec betragen können. Im Vergleich der Einschwingvorgänge des Horns mit und ohne Dämpfer weisen gerade zu Beginn die Spektren große Ähnlichkeiten auf, da sich das ausgedehntere Spektrum des Horns mit Dämpfer erst im Verlauf des Einschwingvorgangs ausbildet.

Zu nicht vernachlässigbaren Unterschieden würde zudem die Existenz von Vorläufergeräuschen oder -tönen für das Horn bzw. die Baßflöte führen. Diese sind für die Identifikation dieser Instrumente bei stärkerem Blasansatz oftmals sehr nützlich. Deshalb wurde in dieser Arbeit jedoch gezielt ein weicher Ansatz für die Stimuli gewählt. Schließlich sind auch bei zehn der Stimuli keine Vorläufergeräusche bzw. -töne in den Einschwingvorgängen nachweisbar, womit ein Einfluß durch sie ausgeschlossen werden kann. Schwache bis mittelstarke Vorläuferimpulse sind lediglich bei den Stimuli der Tonhöhe g für beide Instrumentalvariationen des Horns zu erkennen. In der Interpretation des MDS-Ergebnisses sollte daher am Ende darauf geachtet werden, ob sich diese Stimuli zu den restlichen Stimuli in einer besonderen Form verhalten.

Mit der Vorbetrachtung und Heraushebung von Besonderheiten bezüglich der Gesamtheit der Stimuli konnte bereits ein Ausblick auf zu berücksichtigende bzw. zu erwartende Einflüsse für die MDS gegeben werden. Deutliche Unterschiede zwischen Stimuli lassen sich wie gesehen auf sowohl spektrale als auch temporale Aspekte der Klänge zurückführen. Nach der Betrachtung der Stimuli folgt im nächsten Kapitel zunächst die Einbindung der aufbereiteten Stimuli in die Konzeption und Durchführung des Versuchs.

# 3 - VERSUCHSPLAN & VERSUCHSDURCHFÜHRUNG

In diesem Kapitel wird der Versuch vorgestellt, aus dem die Daten für die MDS erhoben werden. Dies beinhaltet sowohl die Darbietungsform der Stimuli-Paare, ihre Randomisierung und die Versuchsanweisung, als auch die konkrete Versuchsdurchführung, inklusive der technischen Umsetzung. Der Versuch ist grundlegend an dem Versuch von Handel & Erickson (2004) angelehnt. Während Änderungen bezüglich der Instrumente und Tonhöhen im vorangegangenen Kapitel behandelt wurden, wird in diesem Kapitel auf Unterschiede im Versuchsplan und der Versuchsdurchführung eingegangen.

## 3.1 - Versuchsplan

Der Versuchsplan umfaßt die konkrete Gestaltung des Versuchs. Für die MDS dieser Arbeit werden Ähnlichkeitsurteile über Stimuli-Paare benötigt. Diese werden durch die Bewertung von paarweise dargebotenen Stimuli direkt erhoben. Zuvor sollte jedoch geklärt werden, wie sich diese Stimuli-Paare im einzelnen zusammensetzen und darüber hinaus alle Paare dargeboten werden. Außerdem muß für die direkte Bewertung der Ähnlichkeit eine geeignete Ratingskala gewählt werden.

### 3.1.1 - Zusammensetzung und Darbietungsform der Stimuli-Paare

Aus den 12 Hörstimuli ($n$) werden aus allen möglichen Paarungen ($n^2$), abzüglich aller 12 Paare aus identischen Stimuli und die 66 Paare mit inverser Stimuli-Reihenfolge, 66 zu bewertende Paare ($N_{Paare}$) gebildet, entsprechend folgender Formel:

$$N_{Paare} = \frac{n \cdot (n-1)}{2} \qquad \text{Formel 3.1}$$

Somit gibt es sowohl Paare mit Stimuli eines selben Instruments auf unterschiedlichen Tönhöhen, als auch Paare mit Stimuli verschiedener Instrumente bzw. Instrumentalvariationen mit gleicher oder unterschiedlicher Tonhöhe.

Jedes Paar wird zweimal dargeboten und auf Wunsch der VP beliebig oft wiederholt. Zwischen den zwei fünfsekündigen Stimuli eines Paars gibt es eine Pause mit einer Länge von 500 msec. Diese Pausendauer befindet sich Snyder zufolge im Übergangsbereich zwischen dem rein sensorischen Echogedächtnis und dem Kurzzeitgedächtnis.[49] Mit der Ansiedlung der Pausendauer im Anfangsbereich des Kurzzeitgedächtnisses kann der Einfluß einer gezielten, aktiven Gedächtnisleistung beim Vergleichen aufeinanderfolgender Stimuli minimiert werden. Ferner wählten auch Handel & Erickson (2004) diese Dauer als Pause zwischen den Stimuli. Zwischen der ersten und der wiederholten Darbietung eines Paars gibt es eine Pause von 3 sec Länge. Der Versuch gliedert sich in zwei Hälften mit jeweils 33 zu bewertenden Stimuli-Paaren. Zwischen den beiden Teilen gibt es eine kurze Pause.

49 Snyder (2000). S. 19. Laut Snyder dauert das auditive Echogedächtnis (*echoic memory*) in der Regel 250 msec an und kann möglicherweise bis zu einigen Sekunden betragen. Dagegen wird für das Kurzzeit- bzw. Arbeitsgedächtnis eine Dauer von 3 bis 5 sec genannt.

Bezüglich eines für eine MDS dienenden Versuchs mit zu bewertenden Paaren nennen Borg & Groenen zwei Faktoren, durch welche sich Effekte der Reihenfolge auf die Bewertung eines Paars auswirken können.[50] Zum einen gibt es sogenannte *position effects*, bei denen die feste Position eines Stimuli-Paars innerhalb aller Paare des Versuchs zu Ankereffekten führen kann. Außerdem kann durch *timing effects* die Reihenfolge der Stimuli innerhalb eines Paars einen Effekt haben, indem sich die Bewertung der Stimuli-Reihenfolge *A zu B* von der inversen Reihenfolge *B zu A* unterscheidet. In der Konzeption des Versuchsdesigns wird versucht, diesen Effekten durch die angeführten Randomisierungstechniken zu begegnen:

Für die einzelnen Versuchsreihen werden vier randomisierte Reihenfolgen aller Paare generiert, welche auf alle 20 VP gleichmäßig verteilt werden. Somit erhalten je 5 VP eine andere Reihenfolge der Stimuli, wodurch sich der Einfluß von *position effects* verringern läßt. Um ferner den Einfluß der Stimuli-Reihenfolge durch *timing effects* zu begegnen, wird 50 % der VP die inverse Stimuli-Reihenfolge dargeboten.

### 3.1.2 - Bewertungsskalierung

Für jedes Stimuli-Paar wird die Ähnlichkeit in der Klangfarbe direkt bewertet.[51] Die Ähnlichkeiten werden mittels einer diskreten, sechs-stufigen bipolaren Ratingskala bewertet. Die Stufen werden mit Ziffern von ‚1' bis ‚6' versehen. Zudem erhalten die Extremen ‚1' und ‚6' die verbalen Labels *sehr ähnlich* bzw. *sehr unterschiedlich*. Die Labels werden in direkter Anlehnung an Handel & Erickson (2004) gewählt, in denen die englischen Begriffe *very similar* und *very different* verwendet wurden. Abweichend von ihrer Studie, wird zugunsten einer besseren Übersichtlichkeit zudem auf die Verwendung einer acht-stufigen Skala verzichtet. Es sei jedoch angemerkt, daß in MDS-Untersuchungen zur Klangfarbe teilweise auch kontinuierliche Ratingskalen, d.h. mit unbegrenzt vielen, möglichen Unterteilungen, eingesetzt wurden.[52]

Für die Bewertungen wird ein Ordinalskalenniveau angenommen, welche den Anforderungen einer nicht-metrischen bzw. ordinalen MDS genügt. Laut Orth besteht Ungewißheit darüber, welches Skalenniveau durch ein Ratingverfahren erreicht wird. Er führt jedoch weiter aus, daß „in der Regel vermutet wird, daß Ratings [...] aber mindestens zu Ordinalskalenniveau führen"[53]. Problematischer verhält es sich mit der Annahme eines Intervalskalenniveaus, worauf jedoch an dieser Stelle nicht weiter eingegangen werden muß.

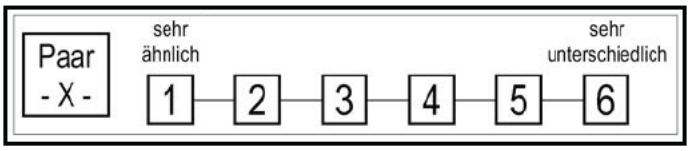


Abb. 3-1

50 Borg & Groenen (2005). S. 114-115.
51 vgl. Sixtl (1982). S. 292-293.
52 vgl. Iverson (1995), Kendall et al. (1998) und Marozeau et al. (2003).
53 Orth (1979). S. 32.

## 3.2 - Versuchsdurchführung

Dem eigentlichen Versuch ging eine Einführung voraus, in welcher die VP mit der Versuchsprozedur vertraut gemacht wurden. In der Einführung wurden den VP alle Stimuli einzeln in Zufallsreihenfolge akustisch dargeboten. Dies diente dem Kennenlernen der Stimuli. Zudem wurden die VP dazu aufgefordert, vorab auf den Umfang bzw. die Bandbreite der verwendeten Klangfarben zu achten. Mit dem vorherigen Vertraut machen mit den Stimuli sollte es den VP ermöglicht bzw. erleichtert werden, gleich von Beginn an die Bewertungsskala richtig einteilen zu können.

Im Anschluß wurden zwei Beispielpaare vorgestellt. Paar 7|2 (Horn normal [g] - Bassflöte [$c^1$]) wurde mit aufsteigender Tonhöhe, Paar 12|4 (Horn mit Dämpfer [$h^1$] - Klarinette [g]) mit absteigender Tönhohe vorgestellt. Dies diente dem Kennenlernen der Versuchs- und Bewertungssituation. Dabei wurden die VP aufgefordert, probeweise auf den Ratingskalen die Ähnlichkeit der Stimuli-Paare zu bewerten. Im eigentlichen Versuch machten die VP ihre Bewertungen auf einem Bewertungsbogen mit je 33 numerierten Bewertungsskalen. Für beide Teile des Versuchs gab es je einen Bewertungsbogen. Als Orientierungshilfe wurde vor der Darbietung jedes Paars eine der Reihenfolge entsprechende Paarnummer genannt. Der gesamte zeitliche Versuchsaufwand umfaßte etwa eine Stunde.

Die VP erhielten bereits vor den zwei Probebewertungen die Versuchsanweisung. Darin wurden die VP aufgefordert nur die Ähnlichkeit in der Klangfarbe zu bewerten und Unterschiede in der Tonhöhe, in der Lautheit, im Vibrato und in der Tondauer außer Acht zu lassen. Darüber hinaus sollten sie bei jeder Bewertung auf eine einheitliche Verwendung der Ratingskala achten. Der genaue Wortlaut ist in Abb. 3-2 nachzulesen. Die Einführung und die Versuchsanweisung wurde den VP sowohl schriftlich als auch mündlich mitgeteilt. Die Versuchsanweisung unterscheidet sich im Vergleich etwas von der von Handel & Erickson (2004), indem bei diesem Versuch der Begriff *Klangfarbe* ausdrücklich genannt wird, während bei ihnen eine Formulierung gewählt wurde, welche seine Verwendung zwar vermeidet, im Resultat jedoch auch unklar bzw. unpräzise bleibt (siehe Abschnitt 1.2.1). Da es sich beim Gegenstand der Forschung letztlich jedoch um die *Klangfarbe* handelt, wurde in dieser Arbeit der direkte Gebrauch des Begriffs gewählt. Es ist jedoch wichtig, diese Abweichungen bzw. auch aus sprachlichen Inkompatibilitäten herrührende semantische Unterschiede für eventuelle Diskrepanzen in den Ergebnissen im Auge zu behalten.

**Einführung**

1) Dieser Versuch dauert etwa eine Stunde. Sie bewerten Paare von Instrumentalklängen hinsichtlich der Ähnlichkeit ihrer Klangfarben auf einer Skala. Infolge werden Ihnen alle in diesem Versuch verwendeten einzelnen Instrumentalklänge zum Kennenlernen vorgestellt. Achten Sie auf die Bandbreite der vorhandenen Klangfarben, um später bei Ihren Bewertungen die Skala entsprechend nutzen zu können.

2) Infolge werden zwei Beispielpaare präsentiert. Jedes Paar wird zweimal präsentiert. Bei Bedarf kann es weitere Male wiederholt werden. Vor der Präsentation eines Paares wird die jeweilige Paarnummer genannt. Bewerten Sie die beiden Paare probeweise in den unten aufgeführten Bewertungsskalen.

**Versuchsanweisung**

Beispiel: Bewertungsskala

| Paar -X- | sehr ähnlich | | | | | sehr unterschiedlich |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |

| Paar -Y- | sehr ähnlich | | | | | sehr unterschiedlich |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |

- Bewerten Sie für alle Klangpaare die Ähnlichkeit der Klangfarben auf einer Skala von 1 bis 6. 1 wird als ‚sehr ähnlich' und 6 als ‚sehr unterschiedlich' gewertet. Kreuzen Sie eine entsprechende Zahl an.
- Achten Sie bei jeder Bewertung auf eine einheitliche Verwendung der 6-stufigen Skala.
- Lassen Sie bei Ihrer Bewertung alle Unterschiede in der Tonhöhe, in der Lautheit, im Vibrato und in der Tondauer außer Acht. Bewerten Sie **nur** die Unterschiede in der Klangfarbe.

Abb. 3-2

### 3.2.1 - Versuchspersonen

Konzipiert wurde der Versuch für 20 VP. Es nahmen 21 Personen an dem Versuch teil. Eine VP (*4f*) wurde letztlich zugunsten von versuchskonzeptionellen Gründe (auf 20 VP ausgelegte Randomisierung), vor allem jedoch wegen Bewertungen mit einem starken Bias, aus der weiteren Betrachtung genommen.[54] Der Bias äußerte sich dadurch, daß über 40 der 66 Bewertungen von dieser VP mit der Extremkategorie ‚6' (*sehr unterschiedlich*) bewertet wurden. Bei einer solch deutlichen Tendenz zu einer Extremkategorie, darf angezweifelt werden, daß allen Paarbewertungen dieser VP ein Ordinalskalenniveau zugrunde liegt.

Alle VP waren musikalisch vorgebildet und waren dabei entweder Studenten der Musikwissenschaft oder Mitglieder eines Sinfonieorchesters. Um eine ausgeprägtere Voreinstellung gegenüber eines der im Versuch verwendeten Instrumente zu verhindern, nahmen keine Personen mit langjähriger Spielpraxis mit diesen Instrumenten (Querflöte, Klarinette, Horn) am Versuch teil.

Von allen 20 VP wurden vorab Informationen über das Alter, das Geschlecht und eventueller Beeinträchtigungen des Hörvermögens erhoben. Das Alter der VP lag zwischen 17 und 30 Jahren. Von den VP waren neun Personen männlich und elf Personen weiblich. Nur bei einer VP (*2b*) lag in der Vergangenheit (d.h. nicht zum Zeitpunkt des Versuchs) ein eingeschränktes Hörvermögen vor, ausgedrückt durch eine Hörabsenkung des rechten Ohrs. Im Vergleich der einzelnen Paarbewertungen von VP *2b* mit den Mittelwerten aller restlichen VP kommt es zu einer starken Korrelation ($r = 0{,}7312$, $p < 0{,}0001$). Zudem sind die Zentraltendenzen von VP *2b* und den anderen VP annähernd gleich. Daher wurden die Daten der VP *2b* für die weitere Betrachtung beibehalten.

54 Die einzelnen Bewertungen der VP sind im Anhang A tabellarisch aufgeführt.

### 3.2.2 - Technische Umsetzung des Versuchs

Als Räumlichkeit für den Versuch diente das kleine Tonstudio des Elektronischen Studios der Technischen Universität Berlin. Dieser Raum bot eine ruhige, geräumige, sowie angenehme Umgebung für die Versuchsdurchführung. An einer Versuchsreihe nahmen bis zu drei Personen gleichzeitig teil.

Die Wiedergabe der Stimuli erfolgte manuell über einen CD-Player (*Sony CDP-XE530*). Das Audiosignal wurde mittels einer digitalen SP/DIF-Schnittstelle an ein Digitalmischpult (*Yamaha 02R*) weitergeleitet. Das Mischpult diente lediglich der Digital-zu-Analog-Wandlung des Signals. Das analoge Ausgabesignal des Mischpults wurde mittels eines mehrkanaligen Kopfhörerverstärkers (*Behringer Powerplay Pro HA4600*) verstärkt. Es wurde für die bis zu drei gleichzeitig betriebenen Kopfhörer dieselbe Verstärkereinheit benutzt, da diese bereits über drei Kopfhörerausgänge verfügte. So war eine identische Verstärkung des Signals für alle Kopfhörer gewährleistet. In jeder Versuchsreihe wurden die gleichen Verstärkereinstellungen beibehalten, so daß die Wiedergabelautstärke, sowie alle anderen Eigenschaften der Signalkette für alle VP konstant blieb. Als Abhörquelle dienten drei baugleiche Studiokopfhörer (*Audio Technica ATH-M40fs*).

Mit der Vorstellung des Versuchs in seiner Konzeption und Durchführung endet die theoretische Vorbereitungsphase des Versuchs, sowie seine praktische Umsetzung. Im folgenden Kapitel schließt sich bereits die Auswertung der erhobenen Daten an.

# 4 - AUSWERTUNG

Mit der Auswertung folgt im eigentlichen Sinne die Durchführung der MDS, mit den erhobenen Ähnlichkeitsurteilen der Klangfarbe als Eingabedaten. Diesbezüglich ist eine MDS jedoch ein sehr umfangreiches deskriptives Statistikverfahren, mit einer Vielzahl wählbarer Berechnungsalgorithmen und darüber hinaus zu bestimmender Optionen. Zwischen der Aufbereitung der Eingabedaten und dem Akzeptieren eines endgültigen Endergebnisses liegen deswegen zahlreiche Arbeitsschritte, welche in diesem Kapitel im einzelnen vorgestellt werden. Vorab ist es jedoch sinnvoll, mit einer Einführung in die Durchführung einer MDS und ihrer Terminologie zu beginnen.

## 4.1 - Multidimensionale Skalierung und ihre Terminologie

An dieser Stelle sollen für die folgende Dokumentation und Interpretation der MDS einige Termini und vereinfachende Ausdrücke zur Beschreibung der MDS vorgestellt werden. Die folgenden Darlegungen stützen sich überwiegend auf Literatur von Kruskal & Wish (1978) und Borg & Groenen (2005).

### 4.1.1 - Eingabe- und Ausgabegrößen einer MDS

Als Dateneingabe für die MDS dieser Arbeit dienen skalierte Ähnlichkeitsurteile über alle möglichen Paarbildungen aus 12 Stimuli, welche in dieser Arbeit bekanntlich 66 Paarbewertungen umfassen. Alle Ähnlichkeitsurteile können in einer Matrix bzw. einer Halbmatrix (wie in Tabelle 4-1 abgebildet) zusammengefaßt werden. Auch stellt diese Matrixdarstellung das Eingabeformat für die im Statistikprogramm SPSS[55] angebotenen MDS-Algorithmen dar.

| | | Stimuli $S_j$ | | | | | |
|---|---|---|---|---|---|---|---|
| | | $S_1$ | $S_2$ | $S_3$ | $S_4$ | $S_5$ | ... |
| Stimuli $S_i$ | $S_1$ | | | | | | |
| | $S_2$ | $p_{2\|1}$ | | | | | |
| | $S_3$ | $p_{3\|1}$ | $p_{3\|2}$ | | | | |
| | $S_4$ | $p_{4\|1}$ | $p_{4\|2}$ | $p_{4\|3}$ | | | |
| | $S_5$ | $p_{5\|1}$ | $p_{5\|2}$ | $p_{5\|3}$ | $p_{5\|4}$ | | |
| | ... | ... | ... | ... | ... | ... | |

Tabelle 4-1

Ähnlichkeitsurteile werden von hier an, entsprechend der weitläufigen Verwendung in der MDS-Literatur, ‚Proximitäten' (*proximities*) genannt, da eine Ähnlichkeit sich auch als eine ‚Nähe' bzw. eben einer ‚Proximität' zwischen zwei Stimuli auffassen läßt. Ein allgemeiner Ausdruck für die Proximität eines beliebigen Paars ij wird fortan mit $p_{ij}$ notiert. Hierbei sind i und j lediglich Indizes für einzelne Stimuli S, also von 1 bis n, bzw. in diesem Fall folgerichtig 1 bis 12. Indizes i und j unterscheiden sich lediglich im Hinblick auf ihre Positionierung in der Matrix, sowie der Reihenfolge der Stimuli-Darbietung innerhalb eines Paars. Hierbei gliedern die Indizes i die Stimuli $S_i$ nach den Zeilen und analog dazu Indizes j die Stimuli $S_j$ nach Spalten. Beispielsweise entspricht die Proximität

55 Statistical Package of the Social Sciences.

von Stimuli $S_3$ und $S_5$ demnach $p_{3|5}$ bzw. bei inverser Reihenfolge der Darbietung $p_{5|3}$. Entsprechend der in Abschnitt 4.2 noch zu untersuchenden Annahme der Matrix-Symmetrie, wird die gesonderte Betrachtung der Proximität $p_{ji}$ jedoch hinfällig. Somit gilt generell $p_{ji} = p_{ij}$, d.h. die Proximitäten $p_{5|3}$ und $p_{3|5}$ werden zusammengefaßt und treten demzufolge in einem gemeinsamen Feld der Halbmatrix auf.

Als Ausgabegröße einer MDS erhält man für jede Proximität $p_{ij}$ eine entsprechende geometrische Distanz $d_{ij}$. Bei der in dieser Arbeit verwendeten ordinalen MDS handelt es sich genauer sogar um euklidische Distanzen. Je nach R-facher Dimensionalität des Ergebnisses ergeben sich, entsprechend des pythagoräischen Theorems zu rechtwinkligen Dreiecken, die $d_{ij}$ aus den Koordinaten $x_{ir}$ und $x_{jr}$ der r-ten Dimension. Die euklidische Distanz entspricht der Hypotenuse des rechtwinkligen Dreiecks, welche sich über R Dimensionen aufspannt.

$$d_{ij}(X) = \sqrt{\sum_r (x_{ir} - x_{jr})^2} \qquad \text{Formel 4.1}$$

Da es sich um euklidische Distanzen $d_{ij}$ handelt, ergibt sich für die Skalierung der Koordinaten bzw. Koordinatenachsen für alle Dimensionen ein Intervallskalenniveau. Die Distanzen $d_{ij}$ können ebenfalls in einer Matrix dargestellt werden (Tabelle 4-2). Dabei verhält sich die Systematik der Anordnung der Felder analog zu den Proximitäten $p_{ij}$.

| Stimuli $S_i$ \ Stimuli $S_j$ | $S_1$ | $S_2$ | $S_3$ | $S_4$ | $S_5$ | ... |
|---|---|---|---|---|---|---|
| $S_1$ | | | | | | |
| $S_2$ | $d_{2\|1}$ | | | | | |
| $S_3$ | $d_{3\|1}$ | $d_{3\|2}$ | | | | |
| $S_4$ | $d_{4\|1}$ | $d_{4\|2}$ | $d_{4\|3}$ | | | |
| $S_5$ | $d_{5\|1}$ | $d_{5\|2}$ | $d_{5\|3}$ | $d_{5\|4}$ | | |
| ... | ... | ... | ... | ... | ... | |

Tabelle 4-2

### 4.1.2 - Der Zusammenhang zwischen Proximitäten $p_{ij}$ und Distanzen $d_{ij}$

Der funktionale Zusammenhang zwischen den Proximitäten $p_{ij}$ als Eingabe- und den Distanzen $d_{ij}$ als Ausgabegröße einer MDS wird durch die Funktion $f(p_{ij})$ bestimmt. Da es sich bei einer MDS jedoch letztlich um ein iteratives Verfahren zur Generierung von geometrischen Stimuli-Konstellationen X handelt, entsprechen die als MDS-Ergebnis resultierenden Distanzen $d_{ij}(X)$ in den seltensten Fällen genau den Funktionswerten der Funktion $f(p_{ij})$. Daher gilt der Zusammenhang:

$$f(p_{ij}) \approx d_{ij}(X) \qquad \text{Formel 4.2}$$

Dieser Ausdruck schließt eine gewisse Variabilität von $d_{ij}(X)$ zu $f(p_{ij})$ ein, welche man auch als einen Fehler (*error*) $e_{ij}$ pro Paar ij ausdrücken kann. Ferner werden die Funktionswerte von $f(p_{ij})$ zugunsten einer leichteren Schreibweise auch Disparitäten $\delta_{ij}$ genannt bzw. als solche notiert.

$$f(p_{ij}) = \delta_{ij} \qquad \text{Formel 4.3}$$

So läßt sich der Fehler $e_{ij}$ in den meisten Fällen durch folgende Formel ausdrücken:

$$e_{ij}^{\ 2} = (\delta_{ij} - d_{ij})^2 \qquad \text{Formel 4.4}$$

Weiterhin kann man den Fehler $e_{ij}$ auch für alle Paare ij zu einem Einzahlwert zusammenfassen: der sogenannten Größe *Stress* (σ). In dieser Größe werden alle Fehler $e_{ij}$ aufsummiert. In den meisten Fällen wird diese Größe zudem auf Basis des Wertebereichs der Distanzen $d_{ij}$ normiert, da sich ansonsten entsprechend der absoluten Größe der Distanzen $d_{ij}$ auch größere bzw. kleinere Stress-Werte ergeben können, ohne daß diese bereits als gut oder schlecht eingeschätzt werden können. Zusätzlich ermöglicht die Normierung eine Vergleichbarkeit mit Daten, welche auf der gleichen Stress-Variante basieren. Von diesen normierten Stress-Größen existieren unterschiedliche Berechnungsvarianten, von denen drei für diese Arbeit relevante hier genannt seien:

*Stress-1*[2]

$$\sigma_1 = \sqrt{\frac{\sum(\delta_{ij} - d_{ij})^2}{\sum d_{ij}^{\ 2}}} \qquad \text{Formel 4.5}$$

*Normed Raw Stress*

$$\sigma_1^{\ 2} = \frac{\sum(\delta_{ij} - d_{ij})^2}{\sum d_{ij}^{\ 2}} \qquad \text{Formel 4.6}$$

*S-Stress*[3]

$$\sigma_S = \sqrt{\frac{\sum(\delta_{ij}^{\ 2} - d_{ij}^{\ 2})^2}{\sum d_{ij}^{\ 4}}} \qquad \text{Formel 4.7}$$

Die Größe Stress spielt in einer MDS eine entscheidende Rolle, da das iterative Verfahren sich nach der Minimierung von Stress richtet. Vorab wird ein Konvergenzkriterium mit einem entsprechenden Wert bestimmt. Unterschreitet die Minderung des Stress-Werts nach einer Iteration den Konvergenzwert, wird die zu diesem Zeitpunkt erreichte geometrische Konstellation X als MDS-Ergebnis gewählt und das Iterationsverfahren beendet. Borg & Groenen empfehlen eine nachträgliche Korrektur des im MDS-Algorithmus ALSCAL zu schwach gewählten Konvergenzwerts von standardmäßig 0,001 auf 0,000001 oder geringer.[58] Dementsprechend wird in dieser Arbeit ein Konvergenzwert von 0,000001 für MDS mit sowohl dem ALSCAL- als auch dem PROXSCAL-Algorithmus verwendet.

### 4.1.3 - Handhabung gleicher Proximitäten $p_{ij}$ in einer ordinalen MDS

Begünstigt durch eine relativ geringe Zahl von Skalenstufen, treten bei ordinalskalierten Daten in der Regel Fälle auf, in denen einige Bewertungen gleiche Werte besitzen. Für eine ordinale MDS

56 Kruskal, Joseph B. (1964). „Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis“. Psychometrika, 29. S. 1-27.
57 Takane, Yoshio, Young, F. W. & De Leeuw, J. (1977). „Nonmetric individual differences multidimensional scaling: An alternating least-squares method with optimal scaling features”. Psychometrika, 42. S. 7-67.
58 Borg & Groenen (2005). S. 551.

existieren zwei unterschiedliche Ansätze für den Umgang mit gleich großen Proximitäten $p_{ij}$ (*tied data*): die sogenannte *primary* und *secondary approach to ties*.[59] Diese Ansätze lassen sich am besten an einem konkreten Beispiel erklären. Der ‚primäre' Ansatz setzt bei einer Übereinstimmung von Proximitäten $p_{3|5} = p_{2|1}$ nicht auch gleichzeitig die Gleichheit der entsprechenden Disparitäten $\delta_{3|5} = \delta_{2|1}$ voraus. Die einzige Bedingung ist, daß wenn der ordinale Zusammenhang $p_{2|3} \leq p_{3|5} = p_{2|1} \leq p_{4|5}$ vorliegt, die Disparitäten lediglich die beiden Bedingungen $\delta_{2|3} \leq \delta_{3|5} \leq \delta_{4|5}$ und $\delta_{2|3} \leq \delta_{2|1} \leq \delta_{4|5}$ erfüllen müssen und somit gleichzeitig auch $\delta_{3|5} \neq \delta_{2|1}$ gelten darf. Dagegen muß bei der *secondary approach to ties* bei $p_{3|5} = p_{2|1}$ gleichzeitig $\delta_{3|5} = \delta_{2|1}$ gelten. Handel & Erickson (2004) wählten für ihre MDS die *primary approach to ties*. Dieser Ansatz führt zu einer logisch plausibleren Behandlung dieser Proximitäten. Bei einer gleich großen Ausprägung ordinalskalierter Größen vermag das Skalenniveau zwar keine weitere Auskunft über Unterschiede zwischen ihnen zu geben, jedoch kann ebensowenig angenommen werden, daß beide Größen gleich groß bzw. stark ausgeprägt sind, da darüber ebenfalls keine Information vorliegt. Von daher erscheint es durchaus sinnvoll, den Disparitäten eine freiere Behandlung zu erlauben; und das gleichzeitig zugunsten einer eventuell besseren geometrischen Konstellation X.

### 4.1.4 - Bewertung der Qualität eines MDS-Ergebnisses

Ein hinreichend geringer Stress-Wert allein gibt letztlich nur Auskunft über die über alle Paare ij zusammengefaßte Güte der Anpassung an die ordinale MDS-Funktion $f(p_{ij})$. Um auch einen Einblick in die Güte der Anpassung für einzelne Paare ij zu erhalten, empfiehlt sich die graphische Darstellung mittels eines sogenannten *Shepard-Diagramms*. Im einem Shepard-Diagramm werden einzelne Distanzen $d_{ij}$ und ihre entsprechenden Disparitäten $\delta_{ij}$ als eine Funktion der Proximitäten $p_{ij}$ graphisch dargestellt (Abb. 4-1). So wird eine differenziertere Betrachtung der Abstände jedes einzelnen $d_{ij}$ zur MDS-Funktion $f(p_{ij})$ bzw. den Disparitäten $\delta_{ij}$ ermöglicht und es können sich ferner im Stress-Wert nicht abbildbare, grundlegende funktionale Abweichungen feststellen bzw. ausschließen lassen.[60]

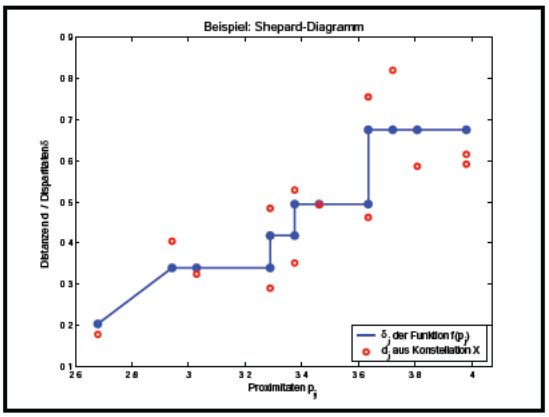


Abb. 4-1

Es kann auch vorkommen, daß die Stimuli $S_i$ in der geometrischen Ergebnis-Konstellation X unterschiedlich gut bzw. angepaßt abgebildet sind. In anderen Worten, ein Stimulus besitzt gegenüber

59 vgl. Borg & Groenen (2005). S. 211-213.
60 vgl. Borg & Groenen (2005). S. 42-44.

einem anderen Stimulus eine in der Konstellation X bessere, d.h. mit geringerer Abweichung zur MDS-Funktion behaftete, Darstellung. Es ist daher ratsam, vor der Interpretation des MDS-Ergebnisses Unterschiede in der Darstellungsqualität einzelner Stimuli-Punkte zu betrachten. Hierzu kann man den sogenannten *Stress-per-Point* (SPP) zu Rate ziehen, dem Stress pro Stimuli-Punkt in X. Hierzu wird für jedes Stimuli $S_i$ der Stresswert zu allen anderen $S_{j \neq i}$ berechnet. Man kann zudem durch 1–SPP den sogenannten *Fit-per-Point* bestimmen und dies in der graphischen dimensionalen Darstellung der Konstellation X mittels größerer bzw. kleinerer Stimuli-Punkte als Parameter darstellen lassen.[61]

Schließlich wird die Größe Stress auch zur Bestimmung der korrekten Dimensionalität einer durchgeführten MDS herangezogen. Hierzu werden MDS für eine Anzahl sinnvoll in Frage kommender Dimensionalitäten durchgeführt. Die hieraus erhaltenen Stress-Werte werden danach als Funktion der Dimensionalität in einem sogenannten *Scree-Diagramm* dargestellt. Man sollte dabei Stress-Werte für ein bis zwei Dimensionalitäten ober- und unterhalb der ‚gewünschten' Dimensionalität berücksichtigen. Generell nimmt der Stress-Wert mit zunehmender Dimensionalität ab, da die Hinzunahme weiterer Dimensionen immer zu einer besseren Anpassung der Distanzen $d_{ij}$ an die MDS-Funktion $f(p_{ij})$ führt. Gleichzeitig kann diese verbesserte Anpassung jedoch auch auf einem sogenannten ‚Fehlerrauschen' beruhen, welches unter anderem aus interindividueller Varianz herrühren kann. Hierbei läuft man Gefahr, eine bestehende Dimension durch eine Vergleichsgröße zu erklären, während sie in Wirklichkeit einzig auf Fehlerrauschen beruht. Als Indiz für die richtige Dimensionalität gilt laut Kruskal & Wish eine Dimensionalität, bei welcher die Stress-Werte der höherdimensionalen MDS-Lösungen keine wesentliche Reduzierung des Stress mehr herbeiführen.[62] Diese Tendenz spiegelt sich (bei Daten mit wenig Fehlerrauschen) durch einen Knick, auch Ellenbogen (*elbow*) genannt, im Scree-Diagramm bei der entsprechenden Dimensionalität wider.

Die endgültige Entscheidung über die den Proximitätsdaten zugrundeliegende Dimensionalität erfordert die Berücksichtigung weiterer Aspekte, auf die an dieser Stelle nicht im einzelnen eingegangen werden kann.[63] In der anschließenden Interpretation des MDS-Ergebnisses werden einige relevante Aspekte jedoch berücksichtigt und zur Ermittlung der endgültigen Dimensionalität eingesetzt.

61 vgl. Borg & Groenen (2005). S. 44-46.
62 vgl. Kruskal & Wish (1978). S. 56.
63 vgl. Kruskal & Wish (1978). S. 53-56 und 89-92, sowie Borg & Groenen (2005). S. 47-55.

## 4.2 - Aufbereitung und Überprüfung der erhobenen Proximitäten

### 4.2.1 - Überprüfung der Matrix-Symmetrie-Voraussetzung

Die MDS dieser Arbeit erfordert 66 Ähnlichkeitsbewertungen entsprechend einer Halbmatrix als Eingabedaten. Wie in Abschnitt 3.1.1 bereits erwähnt, wurde zur Vermeidung von *timing effects* (bezüglich der Stimuli-Reihenfolge innerhalb eines Paars) der Hälfte der VP die inverse Stimuli-Reihenfolge dargeboten. Somit liegen letztlich jedoch Daten einer vollständigen Matrix vor, weil jeweils 10 VP entweder die Proximitäten $p_{ij}$ oder $p_{ji}$ bewerteten. Die Zusammenfassung beider Gruppen von Proximitäten $p_{ij}$ und $p_{ji}$ ist nur unter der Voraussetzung möglich, daß es sich um eine symmetrische Matrix handelt, d.h. alle Bewertungen $p_{ij}$ keinen grundlegenden Unterschied zu den Bewertungen $p_{ji}$ aufweisen.

| $p_{ij}$ vs. $p_{ji}$ | $p_{ij}$ | $p_{ji}$ |
|---|---|---|
| Maximum | 5.40 | 5.30 |
| Minimum | 1.30 | 1.30 |
| arithm. Mittelwert | 3.59 | 3.62 |
| Standardabweichung | 1.19 | 1.11 |

| größte Differenz $\lvert p_{ij}-p_{ji}\rvert$ | 1.2 |
|---|---|

| ij | $p_{ij}$ | ji | $p_{ji}$ |
|---|---|---|---|
| *Maximalwerte* | | | |
| 8\|3 | 4.9 | 3\|8 | 5.4 |
| 12\|1 | 4.7 | 1\|12 | 5.4 |
| 10\|1 | 5.3 | 1\|10 | 4.1 |
| *Minimalwerte* | | | |
| 8\|7 | 1.3 | 7\|8 | 1.3 |
| 2\|1 | 1.3 | 1\|2 | 1.6 |

Tabelle 4-3

Es ergeben sich, wie in Tabelle 4-3 zu sehen, keine gravierenden bzw. systematischen Unterschiede zwischen den Mittelwerten $p_{ij}$ und $p_{ji}$ beider Gruppen[64]. Die Maximalwerte für $p_{ij} = 5{,}4$ und $p_{ji} = 5{,}3$ unterscheiden sich kaum. Zu erwähnen bleibt jedoch, daß unterschiedliche Paare ij und ji für beide Gruppen diesen Maximalwert für das am unterschiedlichsten empfundene Paar erhalten. Außerdem erreichen zwei Paare ($p_{3|8}$ und $p_{1|12}$) den Maximalwert von $p_{ji}$.[65] Die Werte der entsprechenden inversen Reihenfolgen weisen jedoch keinen grundlegenden Unterschied zu den Maximalwerten auf und spiegeln (mit Werten von 4,1 bis 4,9) die Tendenz zu hohen Skalenwerten wider. Für die Minimalwerte ist keine Differenz vorhanden: $p_{ij} = p_{ji} = 1{,}3$. Hier tritt bei $p_{ij}$ für zwei Paare ($p_{8|7}$ und $p_{2|1}$) der gleiche Minimalwert auf. Darüber hinaus bewerteten beide Gruppen dasselbe Paar ($p_{8|7}$ und $p_{7|8}$) als das Paar mit der größten Ähnlichkeit. Der Unterschied für das andere Paar ($p_{2|1}$ und $p_{1|2}$) ist zudem auch sehr geringfügig, womit sich auch hier eine Tendenz zu in diesem Fall kleinen Skalenwerten bestätigt.

Auch die Mittelwerte über alle $p_{ij}$ bzw. $p_{ji}$ weisen mit 3,5894 und 3,6227 nur kleine Unterschiede auf, was ebenfalls für die entsprechenden Standardabweichungen gilt. Die größte absolute Differenz $\lvert p_{ij}-p_{ji}\rvert = 1{,}2$ tritt bei den inversen Paaren $p_{10|1}$ und $p_{1|10}$ auf. Bei demselben Paar tritt auch der Maximalwert für $p_{ij}$ auf. Mit einer maximalen Diskrepanz von 1,2 bei einer Skala von 1 bis 6 fällt der Unterschied jedoch nicht so aus, daß er zu einem gegenteiligen Ergebnis für beide Gruppen

64 Der Begriff *Gruppe* bezieht sich hier nicht auf zwei nach VP getrennte Versuchsgruppen, sondern nur die in dieser Überprüfung benötigte Aufteilung in Proximitäten $p_{ij}$ und $p_{ji}$. Jede VP bewertete dagegen eine gleiche Anzahl Paare ij und ji.

65 Die Zuordnung der Indizes i bzw. j zu den korrespondierenden Stimuli kann anhand Tabelle 4.4 (in 4.2.2) nachvollzogen werden.

führt und ist daher zu vernachlässigen. Ferner sind die gemittelten $p_{ij}$ und $p_{ji}$ beider Gruppen stark positiv korreliert, entsprechend einem Korrelationskoeffizienten von $r = 0{,}8944$ ($p < 0{,}0001$). Somit kann mit ausreichender Sicherheit eine symmetrische Matrix angenommen werden, womit $p_{ij}$ und $p_{ji}$ zusammengefaßt werden können.

### 4.2.2 - Die erhobenen Proximitäten $p_{ij}$

Nach der Erfüllung der Matrix-Symmetrie wird der arithmetische Mittelwert der $p_{ij}$ (bzw. $p_{ji}$) von allen 20 VP gebildet und die Proximitäten $p_{ij}$ in einer Halbmatrix zusammengefaßt. Zuvor wurde überprüft, ob gravierende Unterschiede in den Bewertungsprofilen der VP vorliegen. Für alle Paarbewertungen kommt es zu einer mittleren Standardabweichung über alle VP von 1,07. Die maximale Standardabweichung für eine Paarbewertung erhält man für $p_{10|6}$, mit einem Wert von 1,48; die minimale bei $p_{8|7}$ beträgt dagegen 0,47.[66] Dies deutet nicht auf eine starke interindividuelle Streuung der Paarbewertungen hin. Nur bei einer VP (*2e*) kommt keine signifikante Korrelation ($r = 0{,}1806$, $p = 0{,}1468$) mit dem Mittelwert der $p_{ij}$ der restlichen VP zustande, was jedoch auch bedeutet, daß keine stark negativ korrelierenden Bewertungsprofile (wie durch eine gespiegelte, falsche Verwendung der Ratingskala) vorliegen. Letztlich wird sich für die Beibehaltung und Mittelung der Bewertungen aller 20 VP entschieden.

| S | j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| i | Instrument & Tonhöhe | BFl [g] | BFl [c¹] | BFl [h¹] | Kla [g] | Kla [c¹] | Kla [h¹] | HnN [g] | HnN [c¹] | HnN [h¹] | HnD [g] | HnD [c¹] | HnD [h¹] |
| 1 | Baßflöte [g] | | | | | | | | | | | | |
| 2 | Baßflöte [c¹] | 1.45 | | | | | | | | | | | |
| 3 | Baßflöte [h¹] | 1.75 | 1.95 | | | | | | | | | | |
| 4 | Klarinette [g] | 3.80 | 4.35 | 4.25 | | | | | | | | | |
| 5 | Klarinette [c¹] | 4.10 | 4.40 | 4.60 | 1.55 | | | | | | | | |
| 6 | Klarinette [h¹] | 4.65 | 4.70 | 4.55 | 2.10 | 1.95 | | | | | | | |
| 7 | Horn normal [g] | 4.50 | 4.70 | 5.05 | 3.70 | 3.60 | 3.85 | | | | | | |
| 8 | Horn normal [c¹] | 4.95 | 4.00 | 5.15 | 3.25 | 4.05 | 3.85 | 1.30 | | | | | |
| 9 | Horn normal [h¹] | 4.80 | 4.35 | 4.55 | 4.00 | 4.35 | 4.15 | 2.30 | 1.70 | | | | |
| 10 | Horn mit Dämpfer [g] | 4.70 | 4.90 | 4.60 | 3.60 | 3.40 | 3.75 | 3.00 | 2.85 | 3.50 | | | |
| 11 | Horn mit Dämpfer [c¹] | 4.75 | 4.65 | 4.95 | 3.45 | 3.55 | 4.10 | 1.90 | 2.00 | 2.65 | 2.10 | | |
| 12 | Horn mit Dämpfer [h¹] | 5.05 | 5.10 | 4.50 | 3.65 | 3.65 | 3.50 | 2.20 | 1.90 | 2.30 | 3.30 | 2.15 | |

Tabelle 4-4

In Tabelle 4-4 werden die gemittelten Proximitäten $p_{ij}$ abgebildet.[67] Sie erstrecken sich nahezu über den gesamten Bereich der sechsstufigen Ratingskala von 1 bis 6. Hierbei ist die kleinste Proximität $p_{8|7} = 1{,}3$ und die größte Proximität $p_{8|3} = 5{,}15$. Wie ferner in Abb. 4-2 abgebildetem Histogramm aller $p_{ij}$ zu betrachten ist, weist die Verteilung aller $p_{ij}$ eine Tendenz zu größeren Werten (d.h. größerer Unterschiedlichkeit) auf. Der Median liegt mit einem Wert von 3,825 knapp oberhalb des arithmetischen Mittelwerts von 3,6061, welcher ziemlich genau in die Mitte der Skala fällt. Somit wurden etwa die Hälfte aller Paare als eher unterschiedlich in der Klangfarbe bewertet. Die Ausprägung der

66 Die Standardabweichungen der einzelnen Paarbewertungen werden im Anhang A (in der Tabelle der Einzelbewertungen) aufgeführt.

67 Die in Tabelle 4.4 verwendeten Stimuli-Abkürzungen für die Spalten (j) werden infolge auch in den weiteren graphischen Darstellungen verwendet.

Häufigkeiten am anderen Extrem (d.h. großer Ähnlichkeit) ist weniger stark. Sie liegt unterhalb des Interquartilbereichs und umfaßt somit etwa 25 % aller $p_{ij}$. Es läßt sich zusammenfassend eine Tendenz zu beiden Extremen sowie auch der oberen Hälfte erkennen, während der Bereich direkt unterhalb der Mitte, insbesondere Skalenwerte zwischen 2,3 und 3,3, weitestgehend ausgespart bleibt.

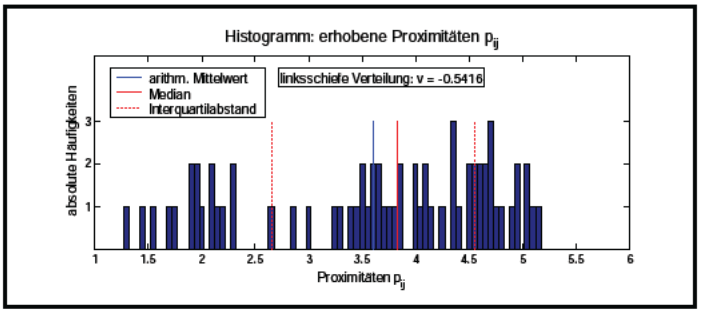


Abb. 4-2

## 4.3 - Durchführung der MDS

Für die konkrete Durchführung der MDS bieten sich unterschiedliche MDS-Algorithmen bzw. -Verfahren an, von denen sich letztlich für eines entschieden werden muß. Innerhalb dieser Verfahren sind zudem unterschiedliche Parameter und Optionen wählbar, welche ebenfalls bestimmt werden müssen. In diesem Teil wird neben der Vorstellung und Besprechung zweier Verfahren und ihren Parametern bzw. Optionen, die Entscheidung für eines von ihnen getroffen, worauf im Anschluß die eigentliche MDS durchgeführt wird. Für die MDS wird eine Anzahl von zwei bis drei Dimensionen als Ergebnis erwartet, womit am Ende auch eine Entscheidung für die zutreffende Dimensionalität erforderlich wird.

### 4.3.1 - Die Algorithmen ALSCAL und PROXSCAL im Vergleich

Das Statistikprogramm SPSS in der Version 12.0 beinhaltet zwei MDS-Verfahren: ALSCAL[68] und PROXSCAL[69] (Version 1.0). Handel & Erickson (2004) führten ihre MDS mit dem ALSCAL-Verfahren durch. Das ALSCAL-Verfahren beruht auf der Minimierung des *S-Stress* (siehe Formel 4.7). Bei *S-Stress* werden entgegen anderer Stress-Größen (*Stress-1* bzw. *Normed Raw Stress*) nicht nur die quadrierten Differenzen zwischen den Distanzen $d_{ij}$ und Disparitäten $\delta_{ij}$ gebildet, sondern $d_{ij}$ und $\delta_{ij}$ werden darüber hinaus selbst einzeln quadriert (siehe Beispiele).

Fehlerberechnung in *S-Stress*: $$e_{ij}^{\ 2} = \left(\delta_{ij}^{\ 2} - d_{ij}^{\ 2}\right)^2$$ Formel 4.8

Fehlerberechnung in *Stress-1*: $$e_{ij}^{\ 2} = \left(\delta_{ij} - d_{ij}\right)^2$$ Formel 4.9

Laut Borg & Groenen werden hierdurch in der geometrischen Konstellation X des Ergebnisses größere Distanzen $d_{ij}$ gegenüber kleineren stärker gewichtet. PROXSCAL hingegen verwendet zur

68 Young, Forrest W., Takane, Y. und Lewyckyj, R. (1978). Thurstone Psychometric Laboratory.
69 Busing, F. M. T. A., Commandeur, J. J. F. und Heiser, W. J. (1997). DTSS, Faculty of Social and Behavioral Sciences, Leiden University.

Stress-Minimierung *Normed Raw Stress*, welche lediglich eine Quadrierung des *Stress-1* darstellt. Im Vergleich zu ALSCAL führt PROXSCAL daher zu einer geringeren Gewichtung größerer Distanzen. Zugunsten einer besseren Abbildung der Feinstruktur innerhalb einer Gruppierung von Stimuli wäre die Nutzung von PROXSCAL daher eventuell von Vorteil.

Als Vorgriff auf das Ergebnis dieser Arbeit erhält man eine Trennung von Stimuli in einzelne Gruppen. Hierbei ist der relative Unterschied der Distanzen $d_{ij}$ der Stimuli zwischen den Gruppen und innerhalb der Gruppen sehr deutlich. Das ALSCAL-Verfahren könnte diesbezüglich diese relativen Unterschiede in den Distanzen eher überbetonen als in einem Ergebnis mittels PROXSCAL. Bei einer probeweisen Durchführung zweidimensionaler MDS mit beiden Algorithmen treten im Hinblick auf die oben genannten Unterschiede für die erhobenen $p_{ij}$ keine nennenswerten Abweichungen zwischen dem ALSCAL- und dem PROXSCAL-Ergebnis in Erscheinung. Sowohl in den Beziehungen zwischen als auch innerhalb der Stimuli-Gruppen sind die Konstellationen X vergleichbar. Hier tritt demzufolge keine Überbewertung größerer Distanzen im Vergleich der beiden Algorithmen ein. Auf eine graphische Abbildung beider Ergebnis-Konstellationen X wird daher verzichtet.

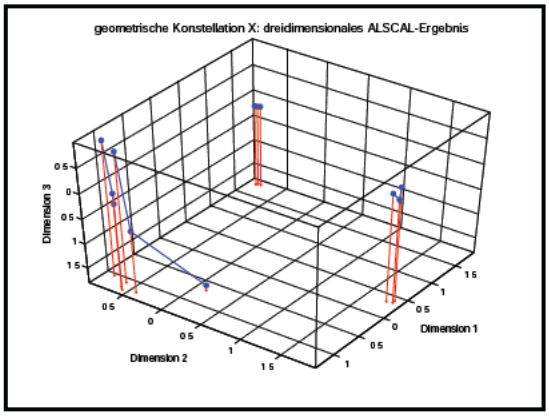


Abb. 4-3

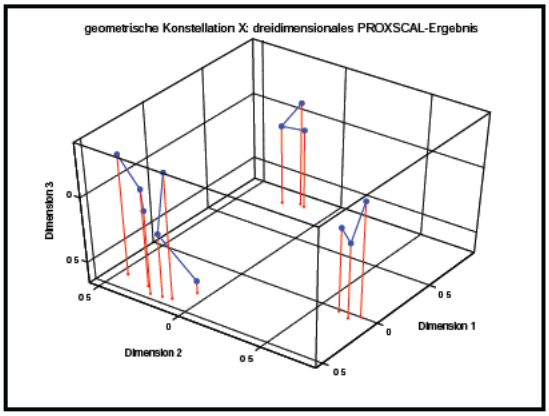


Abb. 4-4

Wie in Abb. 4-3 und 4-4 zu sehen, ist dagegen eine Überbewertung großer Distanzen in der dreidimensionalen Konstellation X der MDS mittels ALSCAL deutlich vorhanden. Während in der dreidimensionalen MDS-Lösung mittels PROXSCAL die Struktur der Stimuli innerhalb der einzelnen Stimuli-Gruppen gut einsehbar ist, tritt die Feinstruktur in der Skalierung von ALSCAL deutlich in den Hintergrund. In der Makrostruktur, d.h. der Positionierung der Stimuli-Gruppen, lassen sich jedoch klare Ähnlichkeiten zwischen den Ergebnissen beider Algorithmen erkennen.[70] Unterschiede in der Orientierung und Skalierung der Dimensionsachsen können ignoriert werden, da es sich bei einer MDS um die Abbildung von relativen Bezügen handelt, bei der Diskrepanzen in den absoluten Skalenwerten nicht entscheidend sind.

Bei Hinzunahme des Shepard-Diagramms (Abb. 4-5) läßt sich für das dreidimensionale ALSCAL-Ergebnis ferner erkennen, daß den Proximitäten $p_{ij} \geq 3{,}5$ (ab der Mitte der Ratingskala aufwärts)

70 An dieser Stelle wird noch nicht auf die inhaltliche Struktur der Stimuli-Verteilung eingegangen. Dieses schließt sich in Kapitel 5 an.

gleichbleibend große Distanzen $d_{ij}$ zugeordnet werden und somit eine Art ‚Kompression' dieses Bereichs auftritt. Somit fallen kleinere $p_{ij}$ gegenüber mittelgroßen und großen $p_{ij}$ in den ihnen zugeordneten Distanzen $d_{ij}$ deutlich kleiner aus. Der PROXSCAL-Algorithmus (Abb. 4-6) läßt dagegen eine deutlich differenziertere Abbildung der $p_{ij}$ in $d_{ij}$ zu, was sich in einem linearen Verlauf der $\delta_{ij}$ ausdrückt.

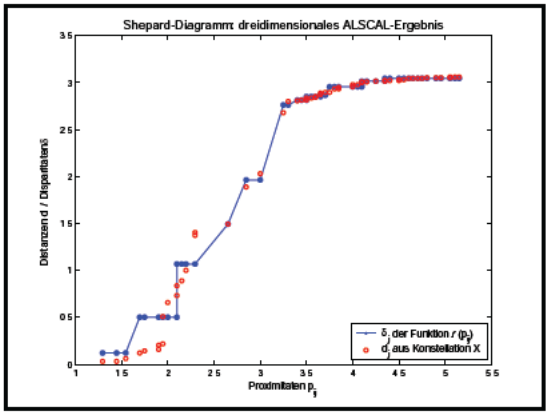


Abb. 4-5

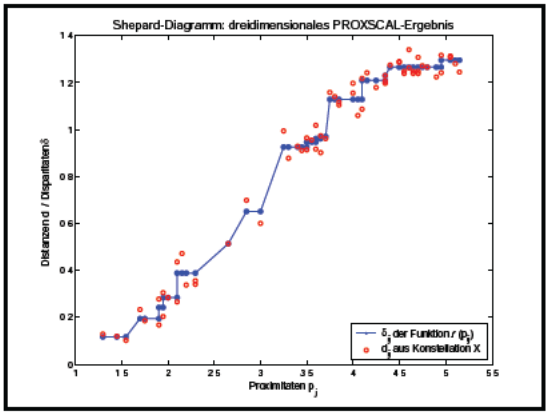


Abb. 4-6

Daher wird für die endgültige MDS für das PROXSCAL-Verfahren votiert, da zudem auch der *Stress-1*-Wert bei diesem Verfahren um etwa die Hälfte kleiner ist (0,03361 für PROXSCAL gegenüber 0,06572 für ALSCAL). Von einer weiteren Berücksichtigung der (durchaus verwendbaren) zweidimensionalen ALSCAL-Lösung wird des weiteren abgesehen, da es bei der späteren Entscheidung zugunsten einer zwei- bzw. dreidimensionalen Lösung nötig ist, daß beide Lösungen auf demselben MDS-Verfahren basieren. Eine Vergleichbarkeit zu den mittels ALSCAL erzielten Ergebnissen von Handel & Erickson (2004) ist Dank der weitestgehend gleichen Makrostruktur der Lösungen beider MDS-Algorithmen weiterhin geboten bzw. anzunehmen.

### 4.3.2 - Die Parameterwahl in PROXSCAL

Das PROXSCAL-Verfahren verfügt neben der zu berechnenden Dimensionalität über weitere wähl- und bestimmbare Parameter bzw. Optionen, welche das MDS-Ergebnis beeinflussen können. In Abschnitt 4.1.2 wurde bereits erwähnt, daß das Konvergenzkriterium für die Stress-Annäherung bzw. -Minderung auf 0,000001 festgelegt wird. Auch wurde sich bereits für die Nutzung des *primary approach to ties* in der Behandlung von gleichen Proximitäten $p_{ij}$ entschieden, welche in der Zuordnung der Distanzen $d_{ij}$ mehr Freiheiten zuläßt.

Als weiterer Parameter läßt sich die Start-Konstellation auswählen, die geometrische Konstellation am Anfang des iterativen Verfahrens. Es stehen hierzu drei Optionen zur Auswahl: *Simplex*, *Torgerson* und *Random*. Laut Borg & Groenen führt eine Torgerson-Start-Konstellation in der Regel zu qualitativ besseren Ergebnissen.[71] In einer MDS existiert zudem immer die Gefahr, bei einer vorliegenden Lösung ein lokales Stress-Minimum als Ergebnis zu erhalten und dieses fälschlicherweise als globales Minimum anzunehmen, da es im unmittelbaren Umkreis zu keiner weiteren Minderung des Stress

71 Borg & Groenen (2005). S. 556.

kommt. Eine Patentlösung zum sicheren Erhalt einer auf dem globalen Stress-Minimum basierenden Ergebnis-Konstellation X existiert nicht. Um die Annahme bzw. den Erhalt eines lokalen Minimums verhindern zu können, existieren jedoch einige Lösungsansätze, auf die wieder nicht im einzelnen eingegangen werden kann. Auch diesbezüglich wird sich auf die Erfahrungen und Empfehlungen von Borg & Groenen gestützt.[72] Hierzu empfehlen sie zur Vermeidung des Erhalts eines lokalen Minimums die *method of dimension reduction.* Statt der alleinigen Durchführung einer MDS in der gewünschten Dimensionalität, werden hierbei weitere MDS für größere Dimensionalitäten durchgeführt. Diese multiplen MDS beginnen bei einer ausgewählten maximalen Dimensionalität. Auf die erste MDS folgen weitere MDS-Durchläufe, in welche sukzessiv jeweils eine Dimension reduziert wird, bis die gewünschte Zieldimensionalität erreicht ist. Entscheidend ist hierbei, daß die in einem MDS-Durchlauf erhaltene Ergebnis-Konstellation X immer als Start-Konstellation des nächsten Durchlaufs verwendet wird.

Nach probeweiser Durchführung von MDS mit PROXSCAL in einigen Parametervarianten stellen sich MDS-Ergebnisse mittels einer sukzessiven Dimensionsreduzierung von 10 bis 2 bzw. 3 Dimensionen und einer Torgerson-Start-Konstellation als am aussagekräftigsten heraus. Darüber hinaus besitzen alle diese mit PROXSCAL berechneten Varianten, wie auch jene mit ALSCAL, in der Makrostruktur der Konstellationen X qualitativ gleichwertige Verteilungen der Stimuli-Gruppen. Deswegen kann mit einiger Sicherheit davon ausgegangen werden, daß das MDS-Ergebnis wirklich auf den Erhalt des globalen Stress-Minimums beruht.

72 Borg & Groenen (2005). S. 276-288.

## 4.4 - Präsentation zwei- und dreidimensionaler PROXSCAL-Ergebnisse

Die endgültige MDS wurde mit der zuletzt genannten Parametervariante durchgeführt. Zur Erstellung eines Scree-Diagramms (Stress pro Dimensionalität) werden zudem noch Stress-Werte (*Normed Raw Stress* und *Stress-1*) für die Dimensionalitäten 1, 4 und 5 berücksichtigt.

### 4.4.1 - Zweidimensionale MDS

Die MDS mit Zieldimensionalität 2 liefert folgende Matrix der Distanzen $d_{ij}$ als Ergebnis:

| S | j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| i | Instrument & Tonhöhe | BFl [g] | BFl [c¹] | BFl [h¹] | Kla [g] | Kla [c¹] | Kla [h¹] | HnN [g] | HnN [c¹] | HnN [h¹] | HnD [g] | HnD [c¹] | HnD [h¹] |
| 1 | Baßflöte [g] | | | | | | | | | | | | |
| 2 | Baßflöte [c¹] | 0.1882 | | | | | | | | | | | |
| 3 | Baßflöte [h¹] | 0.1166 | 0.1011 | | | | | | | | | | |
| 4 | Klarinette [g] | 1.0976 | 1.1610 | 1.1802 | | | | | | | | | |
| 5 | Klarinette [c¹] | 1.2493 | 1.3205 | 1.3359 | 0.1648 | | | | | | | | |
| 6 | Klarinette [h¹] | 1.2919 | 1.3742 | 1.3834 | 0.2439 | 0.0974 | | | | | | | |
| 7 | Horn normal [g] | 1.4101 | 1.3469 | 1.4242 | 0.8759 | 0.9550 | 1.0505 | | | | | | |
| 8 | Horn normal [c¹] | 1.2704 | 1.2022 | 1.2810 | 0.8330 | 0.9342 | 1.0315 | 0.1472 | | | | | |
| 9 | Horn normal [h¹] | 1.3615 | 1.2640 | 1.3527 | 1.0620 | 1.1674 | 1.2647 | 0.2787 | 0.2338 | | | | |
| 10 | Horn mit Dämpfer [g] | 1.4740 | 1.4522 | 1.5133 | 0.6750 | 0.7068 | 0.7944 | 0.3315 | 0.4097 | 0.6073 | | | |
| 11 | Horn mit Dämpfer [c¹] | 1.3917 | 1.3427 | 1.4147 | 0.7798 | 0.8514 | 0.9461 | 0.1096 | 0.1916 | 0.3793 | 0.2285 | | |
| 12 | Horn mit Dämpfer [h¹] | 1.3368 | 1.2628 | 1.3438 | 0.9008 | 0.9962 | 1.0931 | 0.1135 | 0.0762 | 0.1840 | 0.4272 | 0.1989 | |

Tabelle 4-5

Nach Formel 4.1 errechnen sich diese Distanzen $d_{ij}$ aus den zweidimensionalen Koordinaten $x_{ir}$ für alle Stimuli $S_i$. Im folgenden sind die Stimuli-Koordinaten, sowie in Abb. 4-7 ihre geometrische, zweidimensionale Ebenendarstellung, abgebildet:

| Instrument & Tonhöhe | Koordinaten | | Fit-per-Point |
|---|---|---|---|
| | Dimension 1 | Dimension 2 | |
| Baßflöte [g] | 0.9288 | -0.0359 | 0.9966 |
| Baßflöte [c¹] | 0.8827 | -0.2184 | 0.9955 |
| Baßflöte [h¹] | 0.9564 | -0.1492 | 0.9967 |
| Klarinette [g] | -0.0281 | 0.5016 | 0.9912 |
| Klarinette [c¹] | -0.1252 | 0.6347 | 0.9971 |
| Klarinette [h¹] | -0.1108 | 0.7311 | 0.9922 |
| Horn normal [g] | -0.4636 | -0.2584 | 0.9985 |
| Horn normal [c¹] | -0.3180 | -0.2794 | 0.9899 |
| Horn normal [h¹] | -0.3469 | -0.5114 | 0.9978 |
| Horn mit Dämpfer [g] | -0.5418 | 0.0638 | 0.9869 |
| Horn mit Dämpfer [c¹] | -0.4583 | -0.1489 | 0.9979 |
| Horn mit Dämpfer [h¹] | -0.3753 | -0.3296 | 0.9889 |

Tabelle 4-6

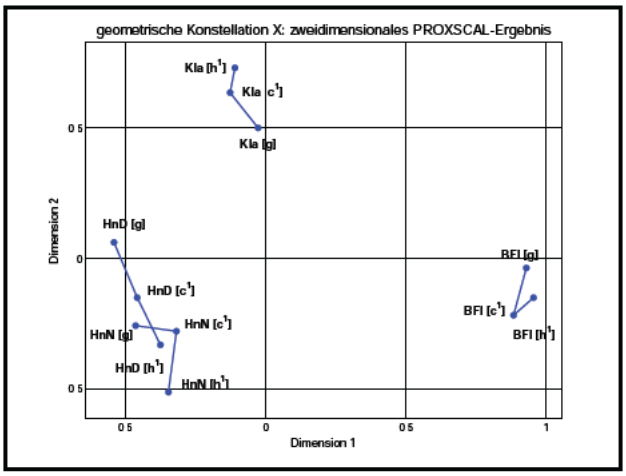


Abb. 4-7

Mit der Größe *Fit-per-Point* (Tabelle 4-6) läßt sich für alle Stimuli $S_i$ die entsprechende Güte der Darstellung in der Konstellation X beschreiben bzw. zeigen. Wie zu sehen ist, besitzen alle Stimuli einen *Fit-per-Point*-Wert von annähernd 1. Daraus läßt sich schließen, daß alle Stimuli in der Konstellation X qualitativ gleichwertig repräsentiert sind. Auf eine in Abschnitt 4.1.4 erwähnte Darstellung der Stimuli als unterschiedlich große Punkte, entsprechend des *Fit-per-Point*-Werts, kann daher wegen der zur erwartenden graphischen Gleichheit verzichtet werden.

Die vorliegende Ergebnis-Konstellation X wurde nach 45 Iterationen erreicht, mit einem *Normed Raw Stress*-Wert von 0,00591. Im folgenden Shepard-Diagramm (Abb. 4-8) läßt sich ferner die Qualität der Anpassung der $d_{ij}$ aus Konstellation X an die MDS-Funktion $f(p_{ij})$ genauer betrachten.

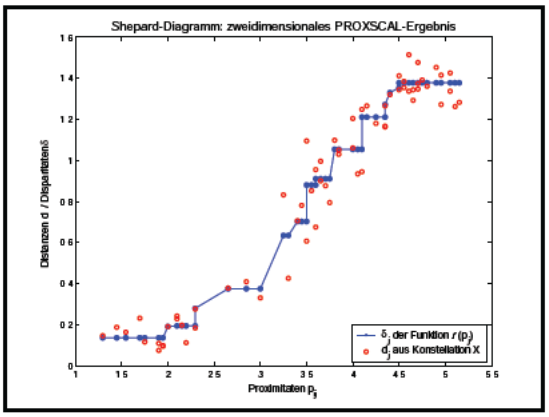


Abb. 4-8

Über alle $p_{ij}$ ist die Abweichung der $d_{ij}$ zu den $\delta_{ij}$ annähernd gleich stark ausgeprägt. Generell läßt sich sagen, daß die $d_{ij}$ dem Verlauf der ordinalen MDS-Funktion bzw. der $\delta_{ij}$ relativ gut wiedergeben. Nennenswert ist die bei kleinen $p_{ij}$ (etwa im Bereich von 1,3 bis 2,3) erkennbare Tendenz in gleich große $d_{ij}$ bzw. $\delta_{ij}$ übertragen zu werden, was sich in einer ‚Kompression' dieses Bereichs ausdrückt. Eine ähnliche Tendenz ist auch bei großen $p_{ij}$-Werten (etwa im Bereich von 4,3 bis 5,3) festzustellen. Dies verursacht einen weniger linearen Verlauf der MDS-Funktion und führt des weiteren zu einer

weniger differenzierten Abbildung der betreffenden $p_{ij}$.

### 4.4.2 - Dreidimensionale MDS

Die MDS mit Zieldimensionalität 3 liefert folgende Matrix der Distanzen $d_{ij}$ als Ergebnis:

| S | j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| i | Instrument & Tonhöhe | BFl [g] | BFl [c¹] | BFl [h¹] | Kla [g] | Kla [c¹] | Kla [h¹] | HnN [g] | HnN [c¹] | HnN [h¹] | HnD [g] | HnD [c¹] | HnD [h¹] |
| 1 | Baßflöte [g] | | | | | | | | | | | | |
| 2 | Baßflöte [c¹] | 0.1190 | | | | | | | | | | | |
| 3 | Baßflöte [h¹] | 0.1883 | 0.2094 | | | | | | | | | | |
| 4 | Klarinette [g] | 1.1419 | 1.1958 | 1.1805 | | | | | | | | | |
| 5 | Klarinette [c¹] | 1.2167 | 1.2731 | 1.2638 | 0.1007 | | | | | | | | |
| 6 | Klarinette [h¹] | 1.2442 | 1.3050 | 1.2479 | 0.2656 | 0.3025 | | | | | | | |
| 7 | Horn normal [g] | 1.2853 | 1.2458 | 1.3045 | 0.9612 | 1.0165 | 1.1032 | | | | | | |
| 8 | Horn normal [c¹] | 1.2405 | 1.1959 | 1.2439 | 0.9935 | 1.0584 | 1.1161 | 0.1267 | | | | | |
| 9 | Horn normal [h¹] | 1.2640 | 1.2070 | 1.2376 | 1.1534 | 1.2285 | 1.2423 | 0.3558 | 0.2347 | | | | |
| 10 | Horn mit Dämpfer [g] | 1.2399 | 1.2205 | 1.3408 | 0.9146 | 0.9297 | 1.1574 | 0.5998 | 0.6971 | 0.9120 | | | |
| 11 | Horn mit Dämpfer [c¹] | 1.2719 | 1.2382 | 1.3145 | 0.9107 | 0.9551 | 1.0859 | 0.1673 | 0.2822 | 0.5119 | 0.4364 | | |
| 12 | Horn mit Dämpfer [h¹] | 1.3096 | 1.2823 | 1.2860 | 0.9017 | 0.9711 | 0.9626 | 0.3399 | 0.2844 | 0.3429 | 0.8792 | 0.4743 | |

Tabelle 4-7

Es folgen die Koordinaten $x_{ir}$ und ihre geometrische Darstellung im dreidimensionalen Raum (Tabelle 4-8 und Abb. 4-9). Zugunsten einer direkten Vergleichbarkeit zur zweidimensionalen Lösung, sowie einer inhaltlich sinnvolleren Anordnung der Stimuli, wurde eine Spiegelung von Dimensionen 2 und 3 (durch Multiplikation mit –1) durchgeführt. Da die Skalierung bzw. Achsenorientierung lediglich der Darstellung von Relationen dient, ist dies problemlos durchführbar und trägt keine Änderung der Konstellation X mit sich.

| Instrument & Tonhöhe | Koordinaten | | | Fit-per-Point |
|---|---|---|---|---|
| | Dimension 1 | Dimension 2 | Dimension 3 | |
| Baßflöte [g] | 0.8255 | 0.1626 | 0.0661 | 0.9998 |
| Baßflöte [c¹] | 0.7884 | 0.2757 | 0.0668 | 0.9990 |
| Baßflöte [h¹] | 0.8403 | 0.1940 | -0.1189 | 0.9988 |
| Klarinette [g] | -0.0161 | -0.6075 | 0.0148 | 0.9989 |
| Klarinette [c¹] | -0.0494 | -0.6828 | 0.0728 | 0.9992 |
| Klarinette [h¹] | 0.0042 | -0.7267 | -0.2217 | 0.9993 |
| Horn normal [g] | -0.4561 | 0.2471 | 0.0165 | 0.9989 |
| Horn normal [c¹] | -0.3981 | 0.3047 | -0.0803 | 0.9980 |
| Horn normal [h¹] | -0.3610 | 0.4603 | -0.2521 | 0.9993 |
| Horn mit Dämpfer [g] | -0.3045 | 0.0624 | 0.5667 | 0.9982 |
| Horn mit Dämpfer [c¹] | -0.4420 | 0.1823 | 0.1702 | 0.9988 |
| Horn mit Dämpfer [h¹] | -0.4311 | 0.1282 | -0.3009 | 0.9982 |

Tabelle 4-8

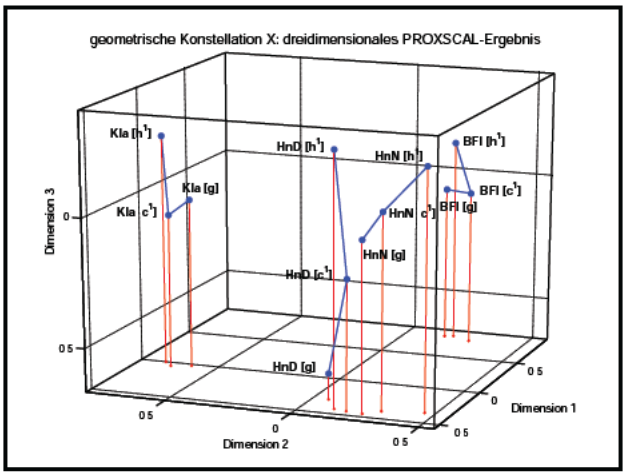


Abb. 4-9

Wie in Tabelle 4-8 zu sehen, sind auch hier die *Fit-per-Point*-Werte aller Stimuli $S_i$ mit annähernd 1 nahezu gleich groß. Somit sind auch hier keine qualitativen Unterschiede in der Güte der Darstellung der Stimuli vorhanden.

Für die dreidimensionale MDS wurde nach 64 Iterationen mit einem *Normed Raw Stress*-Wert von 0,00113 die Ergebnis-Konstellation X erreicht. Im Shepard-Diagramm (Abb. 4-10) läßt sich eine geringere Streuung der $d_{ij}$ um die Funktionswerte $\delta_{ij}$ feststellen als für den zweidimensionalen Fall. Außerdem ist der Stress-Wert um ein Fünftel kleiner als in der zweidimensionalen Lösung. Dies ist jedoch, wie in Abschnitt 4.1.4 bereits besprochen, bei der Hinzunahme weiterer Dimensionen immer zu erwarten.

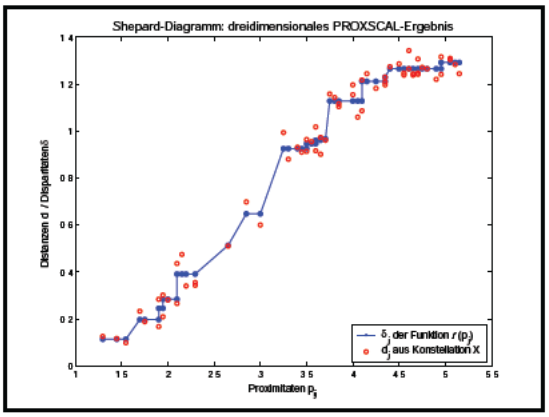


Abb. 4-10

Wie bei der zweidimensionalen Lösung liegt auch hier eine Art ‚Kompression' großer $p_{ij}$ (etwa im Bereich von 4,4 bis 5,0) vor. Jedoch tritt sie hier durch einen insgesamt kleineren Bereich und einer hinzukommenden Abstufung (oberhalb 5,0) abgemildert in Erscheinung. Auch fehlt bei der dreidimensionalen Lösung gänzlich die ‚Kompression' kleiner $p_{ij}$. Im Vergleich zur zweidimensionalen Lösung entspricht die ordinale MDS-Funktion bzw. die $\delta_{ij}$ der dreidimensionalen Lösung daher einem stärker ausgeprägten linearen Verlauf, womit es zu einer besser differenzierten Abbildung aller $p_{ij}$

kommt. Unterschiedliche $p_{ij}$ werden somit, bis auf wenige Ausnahmen, auch durch unterschiedliche $d_{ij}$ ausgedrückt.

### 4.4.3 - Zwei- und dreidimensionale Lösungen im Vergleich

Neben den bisher genannten Ansätzen zur Bestimmung der Dimensionalität empfehlen Kruskal & Wish, basierend auf der Anzahl von Stimuli $S_i$, eine Obergrenze für sicher anzunehmende Dimensionalitäten.[73] Hierfür gilt die auf empirischer Grundlage ermittelte Beziehung:

$$I - 1 \geq 4R \qquad \text{Formel 5.10}$$

*I: Anzahl aller Stimuli $S_i$, R: Anzahl aller Dimensionen r*

Für diese Arbeit mit I = 12 Stimuli würde demnach $11 \geq 4R$ bzw. $2{,}75 \geq R$ gelten. Die Annahme einer Dimensionalität 3 würde diese (Un-)Gleichung damit nicht erfüllen. Andererseits würde jedoch mit R = 3 nicht viel fehlen, um sie erfüllen zu können. Dagegen müßte man von der Annahme einer Dimensionalität R = 4 strengstens abraten.

Wie in Abschnitt 4.1.4 vorgestellt, kann zur Bestimmung der korrekten Dimensionalität auch die Analyse eines Scree-Diagramms hilfreich sein. In Abb. 4-11 dargestelltem Scree-Diagramm sind die Stress-Werte *Normed Raw Stress* und *Stress-1* für die Dimensionalitäten 1 bis 5 graphisch abgebildet. Die in Kruskal & Wish diskutierten Beispiel-Scree-Diagramme beziehen sich lediglich auf *Stress-1*. Zwischen *Normed Raw Stress* und *Stress-1* existiert bekanntlich der simple Zusammenhang einer Quadrierung des Letzteren (siehe Formeln 4.7 und 4.6). Daher ist es sinnvoll, zugunsten einer Vergleichbarkeit zu Kruskal & Wish das Scree-Diagramm für *Stress-1* zu betrachten, da eine Quadrierung zu deutlichen Unterschieden im Funktionsverlauf führt.

| Dimension | Stress-1 | Normed Raw Stress |
|---|---|---|
| 1 | 0.19008 | 0.03613 |
| 2 | 0.07686 | 0.00591 |
| 3 | 0.03361 | 0.00113 |
| 4 | 0.02252 | 0.00051 |
| 5 | 0.01302 | 0.00017 |

Tabelle 4-9

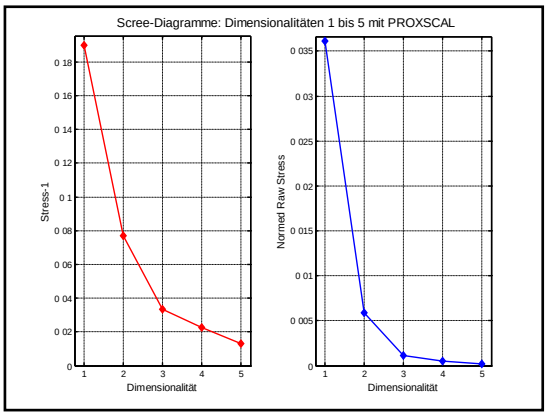


Abb. 4-11

Wie in Abschnitt 4.1.4 bereits erwähnt, deutet im idealen Fall die Lage eines ‚Ellenbogens' die zutreffende Dimensionalität an. Jedoch ist in der Praxis das Vorfinden eines deutlichen Ellenbogens eher unwahrscheinlich, da sich diese nur bei sehr geringem Fehlerrauschen herausbilden. Für den allgemeinen Fall gilt jedoch, daß die zutreffende Dimensionalität dort anzunehmen ist, wo mit der

73 Kruskal & Wish (1978). S. 34.

Hinzunahme weiterer Dimensionen keine weitere wesentliche Reduzierung des Stress eintritt. Dies ist oftmals als ‚Knick' im Verlauf zu erkennen. Im vorliegenden Scree-Plot des *Stress-1* sind zwei solcher ‚Knicke' festzustellen: und zwar bei den Dimensionalitäten 2 und 3. Mit der Hinzunahme einer weiteren Dimension ergibt sich für Dimensionalität 2 eine deutliche Stress-Reduzierung, da von Dimensionalität 2 zu 3 der Stress-Wert mehr als halbiert wird. Dagegen folgt auf Dimensionalität 3 mit den weiteren Dimensionalitäten (mit einer Abnahme um etwa ein Drittel) keine so starke Reduzierung der Stress-Werte. Das Scree-Diagramm deutet daher eher auf das Vorliegen einer Dimensionalität 3.

Handel & Erickson (2004) basieren ihre Feststellungen und Erkenntnisse bekanntlich auf einer dreidimensionalen MDS-Lösung. Darin zeichnet sich vor allem ein starker Einfluß der Tonhöhe in Dimension 1 ab, welcher darüber hinaus eine strukturgebende Funktion in ihrer Ergebnis-Konstellation X einzunehmen scheint. Ein Vorfinden eines Tonhöheneinflußes im MDS-Ergebnis dieser Arbeit ist daher auch wahrscheinlich und zu suchen.

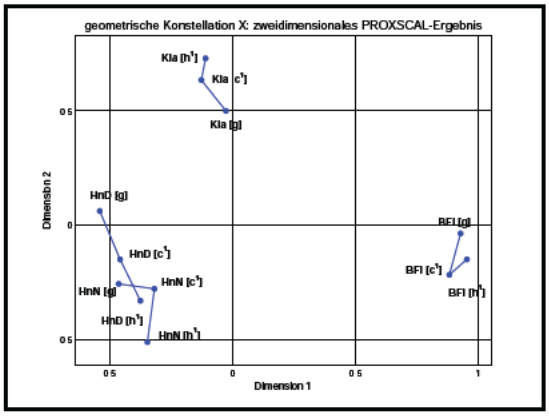


Abb. 4-12

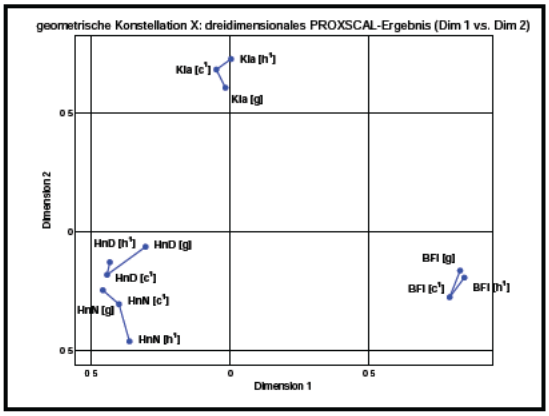


Abb. 4-13

Wie in den graphischen Darstellungen von Dimensionen 1 und 2 der Konstellation X für die zwei- und dreidimensionalen Lösungen zu sehen ist, ist kein ersichtlicher Einfluß der Tonhöhe festzustellen. In der Darstellung der zweidimensionalen Lösung (Abb. 4-12) ist zwar in Dimension 2 eine annähernd lineare Anordnung der Klarinetten- und der Horn-mit-Dämpfer-Stimuli entsprechend der Tonhöhenreihenfolge g–$c^1$–$h^1$ gegeben, nur besitzen diese zueinander eine gegensätzliche Orientierung. Somit ist eine direktional konsistente Interpretation von Dimension 2 als Widerspiegelung eines Tonhöheneinflußes bereits nicht mehr möglich. Außerdem führen diesbezüglich entlang Dimension 2 die keineswegs der Tonhöhenreihenfolge entsprechende Verteilung der übrigen Stimuli zu weiteren Unstimmigkeiten. Auch in der dreidimensionalen Lösung (Abb. 4-13) läßt die sehr ähnlich strukturierte Konstellation in Dimensionen 1 und 2 keinen konsistenten und plausiblen Einfluß der Tonhöhe erkennen.

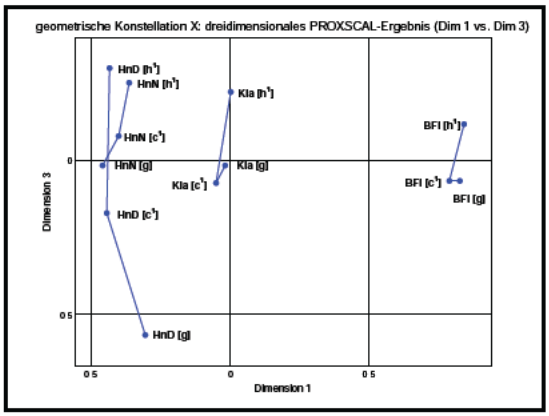


Abb. 4-14

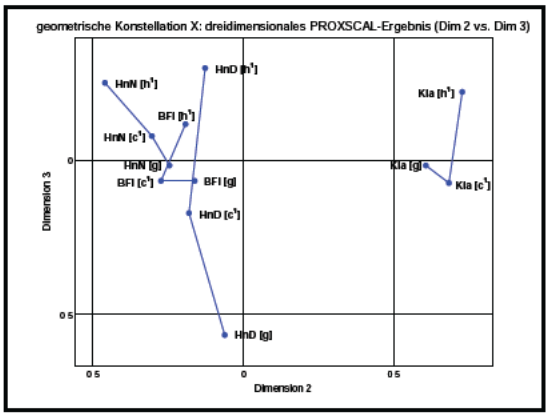


Abb. 4-15

Mit der Hinzunahme einer weiteren Dimension (Abb. 4-14 und 4-15) läßt sich dagegen ein über alle Stimuli bzw. Stimuli-Gruppen konsistenter Einfluß der Tonhöhe in dieser dritten Dimension abbilden. Für jedes Instrument bzw. jede Instrumentalvariation ordnen sich die Stimuli entsprechend der Tonhöhenreihenfolge von negativen zu positiven Skalenwerten von Dimension 3. Ausnahmen bilden hier die Stimuli- bzw. Tonhöhenpaarungen g–$c^1$ der Klarinette und der Baßflöte. Während bei der Baßflöte diese beiden Tonhöhen in Dimension 3 etwa gleich große Skalenwerte besitzen, wird bei der Klarinette sogar die Tonhöhenreihenfolge vertauscht, d.h. die tiefere Tonhöhe g entspricht einem der Tonhöhe $h^1$ näher liegenden Skalenwert als die dazwischen befindliche Tonhöhe $c^1$. Auf mögliche Erklärungen wird an späterer Stelle eingegangen. Allgemein läßt sich jedoch für alle Instrumente bzw. Instrumentalvariationen deutlich eine Tendenz von der tiefsten (g) zur höchsten Tonhöhe ($h^1$) entlang Dimension 3 erkennen. Auch ist der Abstand der beiden tieferen Stimuli-Tonhöhen (g und $c^1$) zur höchsten ($h^1$) ausnahmslos größer als ihre Distanz zueinander. Erwähnenswert ist zudem, daß durch die pro Instrument bzw. Instrumentalvariation unterschiedlich starke Ausbreitung in Dimension 3 eine unterschiedliche Gewichtung des Tonhöheneinflußes für jedes Instrument bzw. Instrumentalvariation zu gelten scheint.

Der in dem MDS-Ergebnis von Handel & Erickson (2004) maßgebliche Einfluß der Tonhöhe läßt sich in dieser Arbeit also nur durch Hinzunahme einer dritten Dimension bzw. der Annahme einer Dimensionalität 3 erreichen. Wie im Scree-Diagramm (Abb. 4-11) zu sehen, ist die Annahme einer dreidimensionalen Lösung hinsichtlich der Stress-Werte durchaus gerechtfertigt. Trotz Nichterfüllung der am Anfang dieses Abschnitts genannten Beziehung von Kruskal & Wish zum Eingrenzen der größt anzunehmenden Dimensionalität, wäre es unter den gegebenen Umständen vertretbar, eine weitere Dimension zur empirisch abgesicherten zweidimensionalen Lösung zuzulassen. Kruskal & Wish meinen diesbezüglich, daß es für die Analyse der Daten oftmals hilfreicher ist, nach einer „geeigneten" (*appropriate*) anstelle der „korrekten" Dimensionalität zu suchen, wobei letztere die hypothetisch, wahre Dimensionalität der Daten darstellt. Dieses gilt vor allem in Hinblick auf die Verwendung einer MDS als deskriptive und explorative Methode um erhobene Daten besser verstehen zu können.[74] Bezüglich einer besseren Interpretierbarkeit der Daten durch Einbeziehung weiterer Dimensionen äußern sich

74 Kruskal & Wish (1978). S. 48.

beide: „In some cases, one may find that the higher dimensions help to clarify the interpretation of the two-dimensional results. When this happens, the structure of the data may be obscured by trying to represent it in too few dimensions.“[75] Hinsichtlich des MDS-Ergebnisses dieser Arbeit ist die Relevanz dieser Feststellung offensichtlich. In der zweidimensionalen geometrischen Darstellung läßt sich kein über alle Stimuli konsistenter Tonhöheneinfluß erkennen. Zwei Instrumente lassen eine Tendenz erkennen, sich entsprechend der Tonhöhenabfolge linear abzubilden, während die Gestalt der restlichen Stimuli-Gruppen diesbezüglich jedoch einige Verwirrung stiftet. Die in Dimensionen 1 und 2 vorherrschende Unklarheit über einen möglicherweise vorhandenen Tonhöheneinfluß läßt sich erst in Dimension 3 auflösen.

Somit spricht einiges – besonders eine bessere Interpretierbarkeit – für die Annahme der dreidimensionalen gegenüber der zweidimensionalen MDS-Lösung. Gleichzeitig sollte man sich aber der knappen Nichterfüllung der Gleichung von Kruskal & Wish bewußt sein. Bei einer Annahme einer höheren Dimensionalität als die durch die Gleichung festgelegte maximale Dimensionalität fordern Kruskal & Wish, die Ergebnisse als vorläufig zu betrachten, bis diese durch Versuche mit einer größeren Anzahl von Stimuli repliziert werden können.[76] Im Gegensatz dazu können Rückschlüsse auf die Bedeutung von Dimensionen 1 und 2 als empirisch abgesichert angenommen werden, da die Strukturen zwischen den zwei- und dreidimensionalen Lösungen nahezu keine Unterschiede aufweisen. Somit ist allein die Bedeutung von Dimension 3 für den geschilderten Einwand wirklich zu berücksichtigen.

75 Kruskal & Wish (1978). S. 58.
76 ebd. S. 34.

# 5 - DEUTUNG DER DIMENSIONEN

Nach der Entscheidung zugunsten der dreidimensionalen MDS-Lösung des PROXSCAL-Algorithmus wird infolge damit begonnen, die geometrische Konstellation X auf vorliegende Gruppierungen der Stimuli $S_i$ bzw. ihrer Verteilung entlang einzelner Dimensionen zu untersuchen, um Rückschlüsse auf die latent wirksamen Einflüsse machen zu können.

## 5.1 - Deskriptive Betrachtung der Konstellation X

Organisationsstrukturen in der geometrischen Konstellation X lassen sich am besten bei gleichzeitiger Darstellung aller Dimensionen erkennen. In der dreidimensionalen graphischen Darstellung der MDS-Lösung (Abb. 5-1) sind alle Stimuli $S_i$ als Punkte entsprechend ihrer in Tabelle 4.8 genannten Koordinaten $x_{ir}$ abgebildet. Zur übersichtlicheren Darstellung der Stimuli eines Instruments bzw. einer Instrumentalvariation werden diese, entsprechend der Tonhöhenreihenfolge g–$c^1$–$h^1$, durch Linien miteinander verbunden. Das erlaubt die vereinfachte graphische Zuordnung von Stimuli eines gleichen Instruments bzw. Instrumentalvariation und erleichtert auch ihre Abgrenzung voneinander. Für eine verbesserte Darstellung ihrer räumlichen Verteilung werden mit Hilfe von roten Linien die Stimuli auf die untere Begrenzungsebene der Koordinatenachsen von Dimension 1 und 2 projiziert.

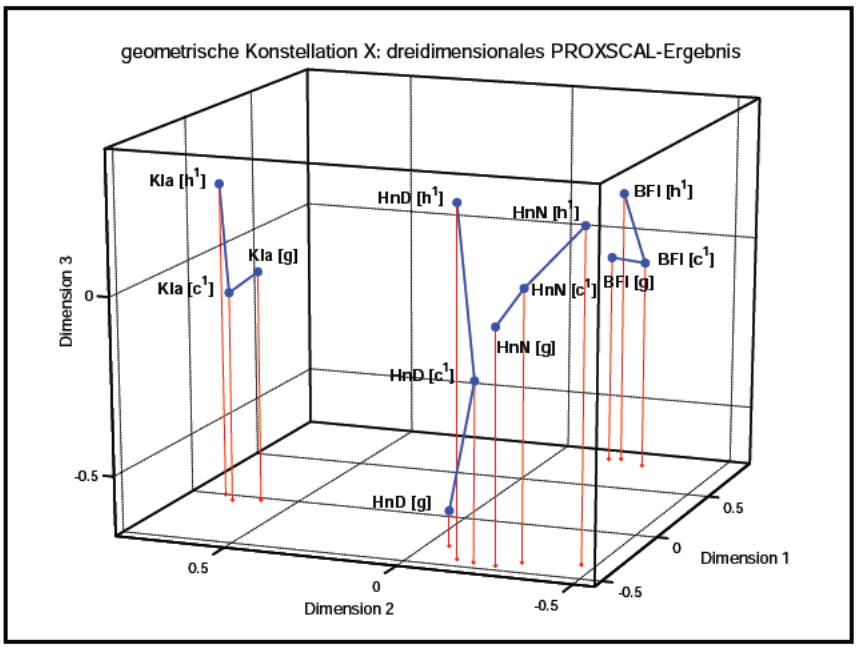


Abb. 5-1

Bezogen auf vorhandene Gruppierungen von Stimuli tritt eine deutliche räumliche Trennung der einzelnen Instrumente auf. Die beiden Instrumentalvariationen des Horns müssen, relativ zu der räumlichen Trennung zwischen den Instrumenten, als eine Gruppe aufgefaßt werden, wobei sie selbst (auf kleinerem Maßstab) eine diskrete räumliche Trennung voneinander aufweisen. Die räumliche Trennung zwischen den Instrumenten läßt sich vor allem auf Unterschiede in den Dimensionen 1 und 2 zurückführen. Dimension 3, welche Anzeichen für einen Tonhöheneinfluß aufweist, trägt nicht zur Trennung der Instrumente bei. Zur Besprechung dieser räumlichen Trennung empfiehlt sich daher eine Beschränkung auf die zwei relevanten Dimensionen.

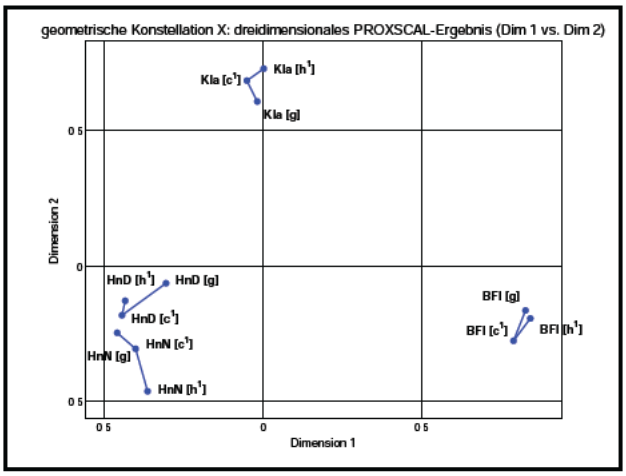


Abb. 5-2

Wie in Abb. 5-2 zu sehen, wird in Dimension 1 die Trennung aller Instrumente erreicht. Beide Varianten des Horns sind in dieser Dimension hingegen nicht voneinander zu unterscheiden. Außerdem befinden sich Klarinette und Horn in dieser Dimension relativ nah beieinander, während sich die Baßflöte deutlicher von der sich am nächsten befindenden Klarinette absetzt. Im Vergleich dazu sind die Beziehungen der Instrumente entlang Dimension 2 weniger klar zu interpretieren. Die Baßflöte und die beiden Varianten des Horns überlappen sich in dieser Dimension. Von ihnen deutlich abgetrennt befinden sich die Klarinetten-Stimuli, womit die Klarinette im vorliegenden MDS-Ergebnis die differenzierteste Abgrenzung zu den anderen Instrumenten aufweist, da diese Abgrenzung auf Unterschiede in zwei Dimensionen beruht. Des weiteren führt Dimension 2 auch zu einer Abgrenzung der Instrumentalvariationen des Horns voneinander, während sich die Baßflöte zwischen ihnen befindet. Diese Feststellung sollte aufgrund des relativ geringen Abstands der Stimuli zueinander jedoch nicht überbewertet werden. Bei zunehmend kleineren Abständen zwischen Stimuli nimmt allgemein die Wahrscheinlichkeit zu, daß zwischen ihnen bestehende Beziehungen verstärkt einer interindividuellen Bewertungsvariabilität unterliegen. Für die Ergebnis-Konstellation X dieser Arbeit bezieht sich dies vor allem auf Stimuli-Beziehungen innerhalb der Instrumente.

Insbesondere gilt dies für die Feinstruktur in Dimension 3, da diese Dimension eben nicht an der Trennung der Instrumente beteiligt ist. In der Makrostruktur zeichnet sich hier ein möglicher Tonhöheneinfluß ab, den man annehmen darf, da erstens sich die Stimuli über größere Abstände von der tiefsten bis zur höchsten Tonhöhe erstrecken und sie sich zweitens in ihrer Orientierung über alle Instrumente bzw. Instrumentalvariationen einheitlich verhalten. Die weniger aussagekräftige Feinstruktur betrifft in dieser Dimension vor allem das Verhältnis der tieferen Tonhöhen g und $c^1$ zueinander. Wie in den Abb. 5-3 und 5-4 abgebildet, verhalten sich diese beiden Tonhöhen über die Instrumente bzw. Instrumentalvariationen gesehen teils unbeeinflußt von (Baßflöte), sowie teils widersprüchlich zu (Klarinette), der in der Makrostruktur vorliegenden Orientierung des Tonhöheneinflusses.

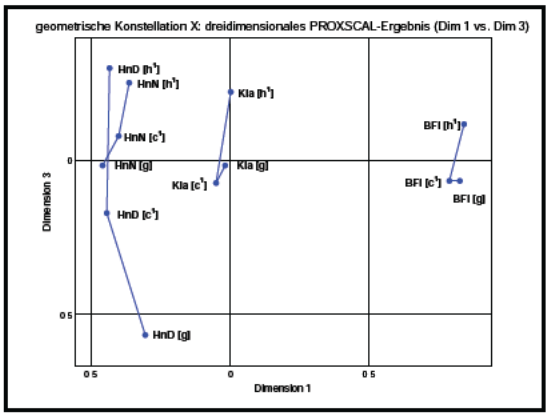


Abb. 5-3

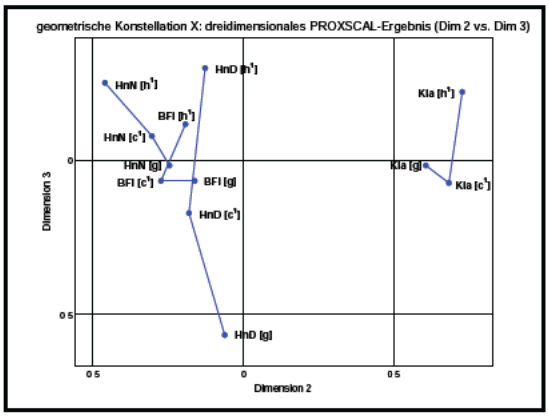


Abb. 5-4

Auffallend ist zudem die pro Instrumentengruppe unterschiedliche Gewichtung der Ausbreitung der Stimuli entlang Dimension 3. Die Stimuli des Horns mit Dämpfer erstrecken sich über einen zwei- bis dreifach größeren Bereich, als die der übrigen Instrumente bzw. Instrumentalvariation. Dagegen ist die Verteilung der übrigen Stimuli in Dimension 3 annähernd vergleichbar. Als Gemeinsamkeit aller Instrumente bzw. Instrumentalvariationen kann erwähnt werden, daß die Paarung g–$c^1$, analog zum einfachen Vergleich der Intervallgrößen (Quarte kleiner als Septime), für alle einem geringeren Abstand entspricht als für die Paarung $c^1$–$h^1$. Zusätzlich läßt sich eine Art ‚Verankerung' aller Instrumente bzw. Instrumentalvariationen entsprechend der höchsten Tonhöhe $h^1$ im Skalenbereich von Dimension 3 (zwischen 0,1 bis 0,3) erkennen. Die bereits erwähnte unterschiedliche Gewichtung in der Skalierung erstreckt sich dahingehend von diesem gemeinsamen Ankerbereich zu negativen Skalenwerten hin. Bezogen auf den angenommenen Einfluß der Tonhöhe für Dimension 3, muß zusätzlich diese instrumentenspezifische Ausbreitung bei der Interpretation des Ergebnisses berücksichtigt werden.

## 5.2 - Ansätze zur Deutung der Dimensionen

Die Folgerung auf die latent wirksamen Faktoren und Einflüsse für die Ergebnis-Konstellation X bzw. den einzelnen Dimensionen wird anhand zweier Ansätze durchgeführt: einmal durch die quantitative Methode des Vergleichs mit objektiven, skalierten Größen, zum anderen durch eine qualitative Methode der hypothetischen und differenzierten Interpretation mit Hilfe von Erkenntnissen aus den Abschnitten 2.4 und 2.5. Beide Verfahren sollen in Ergänzung zueinander gesehen werden, um letztlich zu einer logisch schlüssigen, aber auch empirisch abgesicherten Deutung des MDS-Ergebnisses zu gelangen. Eine alleinige Beschränkung auf den quantitativen Ansatz würde mit den vorhandenen Mitteln nicht quantifizierbare bzw. skalierbare Einflüsse ganz außer Acht lassen. Umgekehrt können durch den qualitativen Ansatz gestellte Hypothesen mit Hilfe quantitativer Befunde bestärkt oder auch entkräftet werden. Da sich diese beiden Ansätze wechselseitig beeinflussen, wäre es wenig sinnvoll sie infolge getrennt zu behandeln. Stattdessen wird eine getrennte, sukzessive Betrachtung der einzelnen drei Dimensionen gewählt.

## 5.3 - Einführende Zusammenfassung der Vergleichsgrößen

Vor der Deutung der Dimensionen ist es sinnvoll, eine Präsentation der im folgenden berücksichtigten objektiven Vergleichsgrößen vorzunehmen, sowie am Ende die Ergebnisse aus den quantifizierbaren Vergleichen mit den Dimensionen tabellarisch zusammenzufassen. Vorab soll zudem spezifiziert werden, auf welcher Grundlage die quantifizierbaren Vergleiche zwischen den skalierten physikalisch-akustischen bzw. psychoakustischen Größen und der dimensionalen Skalierung der Stimuli $S_i$ beruhen.

Alle der einbezogenen Vergleichsgrößen besitzen Intervall- oder sogar Ratioskalenniveau, da sie aus entweder rein physikalischen oder psychometrischen Erhebungsverfahren hervorgehen. Die in dieser Arbeit verwendete ordinale MDS besitzt in der Darstellung jeder Dimension (entsprechend der Koordinatenachsen) ebenfalls Intervallskalenniveau. Bei der Gegenüberstellung der Stimuli-Ausprägungen für eine Vergleichsgröße und eine Dimension handelt es sich somit um einen Vergleich von zwei auf äquidistanten Stufen beruhenden, kontinuierlichen Skalen. Wie auch bei Handel & Erickson (2004) angewendet, wird als quantifizierbarer Vergleich die Berechnung der Korrelation zwischen den Skalenwerten aller Stimuli $S_i$ durchgeführt.[77] Als Wert für die Qualität der Ähnlichkeit bzw. Entsprechung der beiden Skalierungen dient Pearsons Korrelationskoeffizient r (Produkt-Moment-Korrelation). Es kann damit überprüft werden, wie gut sich die Skalenwerte der Stimuli $S_i$ in einer MDS-Dimension durch korrespondierende Skalenwerte aus einer Vergleichsgröße erklären lassen. Hierbei wird der lineare Zusammenhang zwischen ihnen betrachtet. Man erhält demnach eine quantifizierbare Aussage über die Relevanz einer Vergleichsgröße als mögliche Einflußgröße für die Skalierung der Stimuli $S_i$ in einer Dimension in Frage zu kommen.

Die Liste berücksichtigter Vergleichsgrößen gliedert sich in drei Untergruppen:

- Die erste Untergruppe enthält die von Handel & Erickson (2004) berücksichtigten Größen, welche sich auch in zahlreichen anderen MDS-Untersuchungen zur Klangfarbe als signifikante Korrelate herausgestellt haben.
- Die zweite Untergruppe setzt sich aus mehreren standardisierten psychoakustischen Parametern zusammen. Diese entstammen Forschungsergebnissen der Psychoakustik, welche sich vor allem auf den Einsatz reiner Töne bzw. geräuschhafter Klänge als Stimuli beschränken.
- In der dritten Untergruppe befinden sich ‚unerwünschte' Größen, welche aufgrund der im Versuchsplan durchgeführten Standardisierung als Einflußfaktoren eine untergeordnete bis keine Rolle spielen sollten. Hierbei werden Einflüsse durch die stereophone Wiedergabe, sowie auch Unterschiede in der Lautstärke der Stimuli $S_i$ berücksichtigt.

77 Organisationsstrukturen innerhalb eines MDS-Ergebnisses orientieren sich nicht zwangsläufig entlang der gegebenen Koordinatenachsen bzw. Dimensionen. Die vorliegende MDS-Lösung weist jedoch keine Stimuli-Strukturen auf, welche sich nach von den Koordinatenachsen abweichenden Richtungen organisieren. Somit kann auf die Überprüfung von Korrelationen mit von den Koordinatenachsen abweichenden Orientierungen verzichtet werden.

### 5.3.1 - Untergruppe 3

Während unterschiedliche Tonhöhen und Instrumente als beabsichtigte unabhängige Variablen im Versuchsplan integriert wurden, sollten Einflüsse der Lautstärke und einer eventuellen Richtungsabhängigkeit in der stereophonen Stimuli-Wiedergabe systematisch verhindert bzw. minimiert werden. Im Idealfall sollten diese Vergleichsgrößen im MDS-Ergebnis also keine tragende Rolle spielen, weswegen auf diese Untergruppe auch als erstes eingegangen wird. Durch aufnahmetechnische Maßnahmen (siehe Abschnitt 2.3.1), wie auch durch die Standardisierung der Lautstärke (siehe Abschnitt 2.3.2) wurde versucht eine Eingrenzung dieser Einflüsse umzusetzen. Jedoch sind beim subjektiven Hörvergleich der Stimuli gerade bezüglich der stereophonen bzw. räumlichen Ausbreitung wahrnehmbare Unterschiede festzustellen, welche die Ähnlichkeitsbewertungen beeinflußt haben könnten. Daher ist eine Untersuchung dieser Faktoren mittels leicht bestimmbarer objektiver Kriterien unerläßlich, um die Einwirkung eines eventuellen Einflusses feststellen und ferner auf eine Dimension beziehen zu können.

*a) stereophones Hören*

Stereophones Hören ist mit dem Richtungshören bzw. der direktionalen Lokalisation eng verbunden. Für die räumliche Ortung spielen sowohl Laufzeit- als auch Intensitätsunterschiede zwischen dem linken und rechten Ohr des Empfängers eine entscheidende Rolle.[78] Bei den erstgenannten Unterschieden kommt der interaurale Laufzeitunterschied $t_{Diff}$ zum tragen, welcher aus einer Kreuzkorrelationsanalyse der Zeitsignale von beiden Ohrkanälen ermittelt wird. Die zeitliche Auflösung dieser Analyse richtet sich nach der Abtastfrequenz (44,1 kHz) und beträgt etwa 0,023 msec (entsprechend einer Abtastperiode). Laut Dickreiter liegt der ebenmerkliche Zeitunterschied bei etwa 0,03 msec.[79] Demnach führt erst ein Zeitversatz von etwa 0,045 msec (entsprechend zwei Abtastperioden) zu einer eventuell wahrnehmbaren Abweichung vom frontalen Schalleinfall. Eine derart große Zeitdifferenz liegt nur bei einem Stimulus (BFl [$c^1$]) vor.

Intensitätsunterschiede beziehen sich auf Unterschiede in der Signalamplitude zwischen den Hörkanälen. Hierfür werden sowohl Unterschiede in der Verhältnislautheit in Sone ($N_{Diff}$), wie auch Schallpegelunterschiede in dB ($L_{Diff}$) berücksichtigt. Ferner werden für diese beiden Größen sowohl richtungsgebundene (Links-Rechts-Orientierung, mit Suffix: *Dir*) als auch richtungsunabhängige (Betragswert, mit Suffix: *Abs*) Fälle berücksichtigt. Ebenmerkliche Unterschiede in der Verhältnislautheit lassen sich aufgrund der verhältnisskalierten Natur dieser Größe nicht formulieren. Auch für Schallpegelunterschiede existiert (wegen der Frequenzabhängigkeit bei der Wahrnehmung) nur ein grober Richtwert von etwa 1 dB als ebenmerklicher Unterschied. Dieser wird in den Stimuli nur in zwei Fällen annähernd erreicht: BFl [$c^1$] (0,96 dB) und HnD [g] (0,69 dB). Es bietet sich jedoch trotzdem an zu untersuchen, ob sich überhaupt eine Korrelation der Lautstärkeunterschiede mit den Dimensionen herausstellt.

Aus der Stereophonie herrührende Faktoren beschränken sich allerdings nicht nur auf Einflüsse

78 vgl. Dickreiter (1997). S. 117-121, sowie Hellbrück (1993). S. 136-141.
79 Dickreiter (1997). S. 118.

des direktionalen Hörens. Durch instrumentenspezifische Abstrahlcharakteristiken könnten sich beim Empfänger des weiteren Unterschiede in den Zeitfunktionen bzw. Spektren der linken und rechten Kanäle feststellen lassen. Dies wird in dieser Arbeit für zeitliche Unterschiede durch den Maximalwert einer Kreuzkorrelationsanalyse zwischen beiden Zeitsignalen ausgedrückt ($Corr_{Temporal}$). Ferner wird für spektrale Unterschiede eine Überprüfung der interauralen spektralen Ähnlichkeit mit der Berechnung der Korrelation zwischen FFT-Amplitudenwerten beider Hörkanäle vorgenommen ($Corr_{Spectral}$).[80]

*b) Unterschiede in der Lautstärke*

Basierend auf der Verhältnislautheit wurde die Standardisierung der Lautstärke mit einer Genauigkeit von 1,0 Sone realisiert (siehe Abschnitt 2.3.2). In dieser Größenordnung liegen zwischen den Stimuli $S_i$ absolute Abweichungen vom Standardwert (15 Sone) vor. Es sei angemerkt, daß dieser Abstand in einigen Fällen allein schon durch interaurale Lautheitsunterschiede leicht überschritten wird. Es wird angenommen, daß sich für die Hörwahrnehmung diese Lautheitsunterschiede jedoch in einer vernachlässigbaren Größenordnung befinden. Um dies quantitativ zu überprüfen, werden als Vergleichsgrößen sowohl die über beide Kanäle gemittelte Verhältnislautheit $N$ (in Sone) als auch der Schalldruckpegel $L_{SPL}$ (in dB) miteinbezogen. Trotz der Standardisierung nach der Verhältnislautheit in Sone können somit eventuelle auf Schalldruckunterschiede zurückzuführende Einflüsse auch registriert werden.

### 5.3.2 - Untergruppe 1

Es steht außer Frage, daß Frequenzspektren und Einschwingvorgänge für die Klangfarbenwahrnehmung eine maßgebliche Rolle spielen. Frequenzspektren bieten einen sehr reichen Informationsgehalt und können daher in vielerlei Hinsicht beschrieben werden. Dadurch wird die Formulierung einer Größe, welche all diese vielschichtigen Charakteristiken berücksichtigt, praktisch unmöglich gemacht. Es können mit Einzahlgrößen also nur punktuell Spektraleigenschaften zusammengefaßt werden. Einschwingvorgänge lassen sich ebenfalls in vielerlei Hinsicht beschreiben; dies auch im Zusammenspiel mit einer getrennten spektralen Betrachtung einzelner Teiltöne. Die offensichtlichste Eigenschaft des Einschwingvorgangs ist seine Dauer. Ferner ist eine Quantifizierung von Tonhöhenunterschieden für diese Arbeit notwendig. Hierzu lassen sich Grundtöne eines Klangs sowohl aus der Periodizität des Zeitverlaufs, als auch der Lage der Teiltöne im Spektrum bestimmen.

*a) Spektraleigenschaften*

Handel & Erickson (2004) nutzen als objektive Einzahlwerte für die Beschreibung von Spektraleigenschaften zwei Größen: *spectral centroid* und *spectral variability*. Diese zwei Größen haben sich bekanntlich in zahlreichen MDS-Untersuchungen zur Klangfarbe als verläßliche, mit einzelnen Dimensionen stark korrelierende Vergleichsgrößen herausgestellt (siehe Abschnitt 1.2.1).

In der Literatur sind Formeln für die Berechnung des *spectral centroids* im Grundsatz miteinander vergleichbar. In der genauen Umsetzung – gerade bezüglich der Analysefrequenzen und ihren

80 Hierfür dient eine FFT mit 4096 Frequenzlinien, entsprechend einer Frequenzauflösung von 10,8 Hz.

entsprechenden Amplitudenwerten – gibt es jedoch ganz unterschiedliche Ansätze, was zu einiger Unklarheit hinsichtlich ihrer Vergleichbarkeit führt. Handel & Erickson (2004) beziehen ihre *spectral centroid*-Werte aus zuvor von Sandell kalkulierten Werten für Klänge aus der Instrumentalklänge-Bibliothek MUMS. In einer Studie im gleichen Jahr (wie die Berechnung der Klänge aus MUMS) bezieht Sandell für die *spectral centroid*-Berechnung benötigte Frequenz- und Amplitudenwerte aus Erregungsmustern der Basilarmembran.[81] Diese Erregungsmuster entsprechen einer psychoakustisch wirksamen Spektraldarstellung, welche der Frequenz- und Amplitudenauflösung der menschlichen Hörwahrnehmung entspricht.

Dagegen stützen sich die meisten Varianten des *spectral centroid* auf Frequenz- und Amplitudenwerte eines mittels FFT ermittelten rein physikalisch-akustischen Spektrums, oder wahlweise auch auf nur jenen Frequenzbändern, in denen sich Teiltöne befinden.[82] Zusätzlich schlagen Kendall et al. eine Normierung nach der Grundtonfrequenz vor[83], welche sich gerade für den repräsentativen Vergleich von Stimuli unterschiedlicher Tonhöhe anbieten würde, da hierdurch eine auf absoluten Frequenzen basierende Abhängigkeit aufgehoben wird.

Aus praktischen Gründen wird zugunsten einer Vergleichbarkeit mit Handel & Erickson (2004) *spectral centroid* (sc) nach Sandell sowie auch mit FFT-Werten berechnet. Für beide Varianten wird zusätzlich noch eine Grundtonfrequenz-Normierung ($f_0$) entsprechend dem Ansatz von Kendall et al. durchgeführt (*Sandell–*$f_0$ bzw. *FFT–*$f_0$), womit letztlich vier unterschiedliche Varianten des *spectral centroid* berücksichtigt werden. Die Berechnungsformeln dieser Varianten werden im einzelnen in Formeln 5.1 bis 5.4 wiedergegeben:

$$sc_{Sandell} = \frac{\sum_{i=1}^{330} f_i \cdot a_i}{\sum_{i=1}^{330} a_i}$$

Formel 5.1

$$sc_{Sandell-f_0} = \frac{\sum_{i=1}^{330} f_i \cdot a_i}{f_0 \cdot \sum_{i=1}^{330} a_i}$$

Formel 5.2

$$sc_{FFT} = \frac{\sum_{i=1}^{512} f_i \cdot a_i}{\sum_{i=1}^{512} a_i}$$

Formel 5.3

$$sc_{FFT-f_0} = \frac{\sum_{i=1}^{512} f_i \cdot a_i}{f_0 \cdot \sum_{i=1}^{512} a_i}$$

Formel 5.4

*i: Frequenzbänder,* $f_i$*: Frequenz des Bandes i,* $a_i$*: Amplitude des Bandes i,* $f_0$*: Grundtonfrequenz*

Den Formeln kann entnommen werden, daß es sich bei dieser Vergleichsgröße um den *Mittelpunkt* bzw. sogar den *Mittelwert* einer Spektralverteilung handelt.[84] Im nicht-normierten Fall erhält man eine Frequenz in Hz als Wert, im normierten Fall dagegen ein einheitenloses Frequenzverhältnis. Für die

81 Sandell (1995). S. 219-222.
82 Es ist nicht klar, ob die von Handel & Erickson (2004) verwendeten *spectral centroid*-Werte wirklich nach der Methode in Sandell (1995) berechnet wurden. Dies läßt sich des weiteren durch deutlich abweichende Wertebereiche zwischen ihren und den nach Sandell (1995) berechneten Werten dieser Arbeit in Frage stellen. Mit der Berücksichtigung von vier Varianten werden aber genügend Optionen geboten, daß in mindestens einem Fall eine Vergleichbarkeit angenommen werden kann.
83 Kendall et al. (1998). S. 348.
84 vgl. Kendall et al. (1998). S. 330, bzw. Iverson & Krumhansl (1993). S. 2595.

$sc_{Sandell}$-Werte wurden 330 Frequenzbänder entsprechend der *equivalent rectangular bandwidth* (ERB)[85] berücksichtigt und die Erregungspegel bei diesen Frequenzen als Amplitudenwerte genommen. Für die $sc_{FFT}$-Werte dient eine FFT-Analyse mit 1024 Frequenzlinien entsprechend einer Frequenzauflösung von etwa 43 Hz. Für den Nutzfrequenzbereich steht jedoch nur die Hälfte der Frequenzlinien zur Verfügung. Entsprechend Handel & Erickson (2004), bzw. den ursprünglich von Sandell berechneten Werten, wird *spectral centroid* erst nach den ersten 500 msec berechnet, um starke Variabilitäten im Einschwingvorgang zu ignorieren. Der letztlich als Vergleichsgröße dienende Einzahlwert ergibt sich aus dem Mittelwert aller pro Zeitfenster berechneten sc-Werte. Die Zeitauflösung entspricht der infolge genannten kürzeren Dauer zur Berechnung der *spectral variability*.

Die Größe *spectral variability* beschreibt die zeitliche Variabilität des Spektrums. Bei Handel & Erickson (2004) entspricht diese Größe der Standardabweichung des zeitlichen Verlaufs des *spectral centroids*. Es werden hierfür aufeinanderfolgende, nicht überlappende Zeitfenster mit zwei unterschiedlichen Zeitauslösungen berücksichtigt (85 und 170 msec).[86] Für die Vergleiche werden die beiden nicht-normierten Fälle des *spectral centroids* berücksichtigt.[87]

*b) Tonhöheneinfluß*

Der starke Einfluß der Tonhöhe bei Handel & Erickson (2004) läßt sich über eine Skala von Halbtönen, als Vergleichsgröße für Tonhöhenunterschiede, quantitativ erfassen. In dieser Arbeit werden Tonhöhenunterschiede in *Cents*[88] ausgedrückt, welche aus Verhältnissen zwischen Grundtonfrequenzen bestimmt werden. Zudem werden auch die Frequenzverhältnisse selbst als Vergleichsgröße berücksichtigt ($f/f_R$). Als Referenzfrequenz bzw. -tonhöhe für die relativen Tonhöhenunterschiede wird die tiefste in den Stimuli vorliegende Grundtonfrequenz von Stimulus HnD [g] gewählt.[89] Die Grundtonfrequenz wird rechnerisch aus der Periodizität der Maxima einer Autokorrelationsanalyse der Stimuli ermittelt. Die Tonhöhenunterschiede $\Delta p$ in Cents werden mit folgender Formel berechnet:

$$\Delta p = \log_2\left(\frac{f}{f_R}\right)\cdot 1200 \qquad \text{Formel 5.5}$$

*f: Grundtonfrequenz; $f_R$: Referenzfrequenz, tiefste vorhandene Grundtonfrequenz*

85 Die ERB-Skala ist mit der Bark-Skala vergleichbar, auf die in Abschnitt 5.4.2 kurz eingegangen wird. Erstere wird bei der Berechnung der Lautheit nach Moore & Glassberg vorgezogen. vgl. Moore, B. C. J. und Glassberg, B. R. (1996). „A revision of Zwicker's loudness model". *Acta Acustica*, 82. S. 335-345.

86 Handel & Erickson (2004). S. 594.

87 Alle Werte für *spectral centroid* und *spectral variability* werden mit *MATLAB* berechnet. Die ERB-Werte basieren auf einer Berechnungsroutine von *EarWig Software* (University of Dublin), abgerufen am 2.6.2006 unter http://www.mee.tcd.ie/mmt/ewd/ewdsoft.php.

88 Die *Cent*-Skala ist eine metrische Unterteilung der Halbton-Skala. 100 Cents entsprechen 1 Halbton.

89 Zwischen Stimuli einer gleichen Tonhöhe gibt es bezüglich der Grundtonfrequenz geringe Unterschiede. Die Abweichungen unterschreiten mit einer Ausnahme jedoch die Unterschiedsschwelle von 3,5 Hz für Frequenzen unterhalb 500 Hz. (vgl. Dickreiter (1997). S. 114.) Die höchste Tonhöhe ($h^1$) besitzt eine Grundtonfrequenz von ca. 495,5 Hz. Die Ausnahme bildet der Stimulus Kla [$h^1$] mit 501,1 Hz.

*c) Einschwingeigenschaften*

Bezüglich der Einschwingvorgänge beziehen Handel & Erickson (2004) als Einzahlwert lediglich die Einschwingdauer ($Onset_{RMS}$) mit ein. Sie wird in ihrem Fall am graphisch dargestellten Verlauf des Effektivwerts (*root-mean-square*) des Signals bestimmt. Die Dauer umfaßt das erste Hervortreten einer periodischen Schwingung aus dem Grundrauschen bis zum Erreichen einer Steigung von annähernd Null.[90] Für die Stimuli $S_i$ dieser Arbeit treten bei der Bestimmung der Dauer nach dieser Methode einige Unklarheiten auf. Während sich für die Stimuli der Klarinette und Baßflöte nach diesem Verfahren das Ende des Einschwingvorgangs eindeutig identifizieren läßt, ist die oben genannte Bedingung bei allen Horn-Stimuli an mehreren Stellen erfüllt. Wie exemplarisch für alle Horn-Stimuli in Abb. 5-5 dargestellt, können drei Stellen als Alternativen für das Ende des Einschwingvorgangs bestimmt werden.

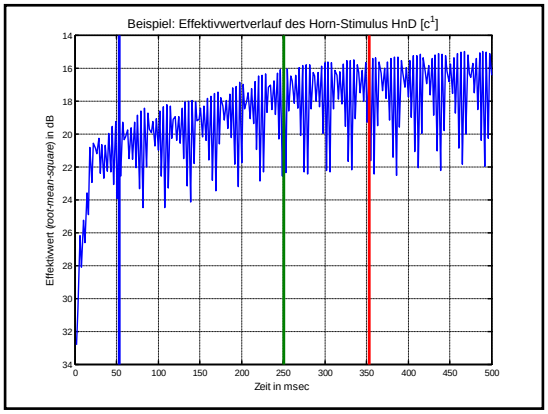


Abb. 5-5

Nach subjektiven Hörvergleich kann die erste Alternative (blau) mit einiger Sicherheit ausgeschlossen werden, da diese Dauer sogar noch kürzer als die der Klarinetten-Stimuli wäre, was sich nicht bewahrheitet. Bei der dritten Alternative (rot) ist keine weitere Zunahme des Effektivwerts zu verzeichnen, was jedoch auch keine zwingende Voraussetzung für das Erreichen des quasi-stationären Schwingungszustands ist. Diese Alternative führt darüber hinaus zu einer insgesamt unrealistischen wirksamen Einschwingdauer (von 250 msec). Somit wird sich für die zweite Alternative (grün) entschieden, da diese in ihrer Größenordnung auch mit den Einschwingdauern aus der Betrachtung in Abschnitt 2.4.2 übereinstimmt.

Für die Relation der Einschwingvorgänge einzelner Teiltöne existieren in der Literatur keine quantitativen Vergleichsgrößen. Daher werden sie bei Bedarf lediglich in der qualitativen Untersuchung berücksichtigt.

### 5.3.3 - Untergruppe 2

Als Teilbereich der Psychophysik wird in der Psychoakustik durch psychometrische Verfahren versucht, einen funktionalen Zusammenhang zwischen objektiv definierbaren, physikalischen Reizen und den durch sie induzierten subjektiven Empfindungen auszudrücken. Aus dem

90 Handel & Erickson (2004). S. 594.

ermittelten funktionalen Zusammenhang läßt sich infolge, nach Bestimmung eines Referenzwerts, eine quantitative Empfindungsgröße formulieren. Alle in dieser Arbeit berücksichtigten psychoakustischen Vergleichsgrößen entstammen der Tradition der Verhältnisskalierung. Unter dieser Skalierungsmethode versteht man eine Skalierung, welche in der Regel auf dem Erhebungsverfahren der *magnitude estimation* (Größenschätzung) beruht. In diesem Verfahren werden empfundene Verhältnisse mittels einer direkten subjektiven Bewertung aus Vergleichen zwischen physikalischen Reizen ermittelt. Der funktionale Zusammenhang läßt sich für verhältnisskalierte Größen oftmals durch eine Potenzfunktion ausdrücken.[91] Wie bereits der Name andeutet, wird für alle durch diese Skalierungsmethode bestimmten Größen Ratioskalenniveau angenommen. Somit eignen sie sich auch für die Berechnung von Korrelationen zwischen Skalenwerten.

*a) psychoakustische Vergleichsgrößen*

Als psychoakustische Vergleichsgrößen dienen in dieser Arbeit die Verhältnisskalen der *Schärfe* (nach Aures), *Rauigkeit*, *Schwankungsstärke* und *Tonhaltigkeit*.[92] Auch *Verhältnislautheit* in Sone ist eine verhältnisskalierte, psychoakustische Größe. Sie fließt ebenfalls in die Berechnung einiger der oben genannten Größen ein. Diese Größen werden für alle Stimuli $S_i$ mit dem Programm *ArtemiS* berechnet. Neben den Einzahlwerten wird zusätzlich der zeitliche Verlauf dieser Größen begutachtet. Die Analyse dieser Zeitverläufe kann jedoch nur rein qualitativ durchgeführt werden und wird deswegen, sofern relevant, erst in der weiteren Besprechung einbezogen werden. Mit der Berücksichtigung der zeitlichen Verläufe kann beispielsweise verhindert werden, daß in den Zeitverläufen vorhandene systematische Unterschiede zwischen Stimuli bei alleiniger Betrachtung der zeitlich gemittelten Einzahlwerte übersehen werden.

Wie eingangs erwähnt, wurden in der psychoakustischen Forschung vor allem synthetische Klänge als Stimuli verwendet. Zum einen waren dies reine Sinustöne bzw. auch Tonkomplexe mit einer geringen Anzahl von Teiltönen. Andererseits wurden viele Versuche auch mit (statistisch-gleichverteilten) Rauschsignalen in unterschiedlichen Varianten durchgeführt, welche einen ‚geräuschhaften' Klangcharakter besitzen.[93] Insofern eignen sich die genannten psychoakustischen Größen zur Beschreibung ‚geräuschbehafteter' Klänge bzw. Geräusche. Sie werden in der Praxis für die Bewertung maschinell-erzeugter Geräusche oder auch ‚Geräuschkulissen' (*soundscapes*) eingesetzt. Geräuschanteile in Instrumentalklängen treten dagegen als nebensächliche Faktoren in Erscheinung. Klangeindrücke von Instrumentalklängen sind vor allem geprägt durch das Vorhandensein deutlich tonaler Komponenten, d.h. im weitesten Sinne einer harmonischen Obertonstruktur entsprechenden Spektralcharakteristik. Daher ist es unklar, ob und inwiefern sich psychoakustische Parameter zur Beschreibung von Instrumentalklängen eignen, da ihre psychometrischen Funktionen auf empirischen Erhebungen beruhen, in denen mit Instrumentalklängen kaum vergleichbare Stimuli eingesetzt wurden. Bezüglich der auf den genannten Größen basierenden

91 Als weiterführende Literatur empfiehlt sich Hellbrück (1993).

92 Für eine Vielzahl der Vergleichsgrößen kann in dieser einführenden Zusammenfassung nicht näher auf ihre Berechnungsmethode, sowie der ihnen zugrundeliegenden Theorie eingegangen werden. Für Vergleichsgrößen, welche zu statistisch signifikanten Korrelationen mit den MDS-Dimensionen führen, wird dies in Abschnitt 5.4 nachgeholt.

93 Eine umfassende Darstellung von Forschungsergebnissen der Psychoakustik liefern Zwicker & Fastl (1999).

psychoakustischen Größe *sensorischer Wohlklang* rät Reuter zur Vorsicht, sollte sie für die Beschreibung von Instrumentalklängen genutzt werden: „[D]as Maß für den sensorischen Wohlklang eignet sich [...] eher für die Beschreibung von Motoren- und Maschinengeräuschen denn als Maßstab für musikalische Wahrnehmungen.“[94]

### 5.3.4 - Korrelationen zwischen den Vergleichsgrößen und Dimensionen

Es folgen in tabellarischer Form die berechneten Korrelationskoeffizienten r (einschließlich der Irrtumswahrscheinlichkeiten p) der Vergleiche zwischen allen Vergleichsgrößen und den drei Dimensionen der MDS-Lösung.[95]

| Vergleichsgrößen | Dimension 1 | | Dimension 2 | | Dimension 3 | |
|---|---|---|---|---|---|---|
| | r | p | r | p | r | p |
| ***Untergruppe 1*** | | | | | | |
| $SC_{Sandell}$ | -0.5756 | 0.0502 | 0.4883 | 0.1072 | -0.0908 | 0.7789 |
| $SC_{Sandell\text{-}f0}$ | -0.4508 | 0.1414 | 0.3056 | 0.3341 | -0.7321 | 0.0068 |
| $SC_{FFT}$ | 0.2950 | 0.3520 | 0.2451 | 0.4426 | 0.5191 | 0.0837 |
| $SC_{FFT\text{-}f0}$ | 0.1037 | 0.7485 | 0.1278 | 0.6924 | -0.7162 | 0.0088 |
| $SV_{Sandell\text{-}85}$ | -0.2008 | 0.5316 | -0.0565 | 0.8616 | -0.6508 | 0.0219 |
| $SV_{Sandell\text{-}170}$ | -0.2033 | 0.5262 | -0.0656 | 0.8394 | -0.6392 | 0.0252 |
| $SV_{FFT\text{-}85}$ | 0.1031 | 0.7499 | -0.3560 | 0.2561 | 0.6202 | 0.0314 |
| $SV_{FFT\text{-}170}$ | 0.1109 | 0.7315 | -0.2892 | 0.3620 | 0.6544 | 0.0210 |
| Frequenzverhältnis $f/f_R$ | 0.0141 | 0.9652 | -0.0303 | 0.9256 | 0.7423 | 0.0057 |
| Tonhöhenverhältnis Δp | 0.0121 | 0.9702 | -0.0375 | 0.9078 | 0.7449 | 0.0054 |
| $Onset_{RMS}$ | -0.4014 | 0.1959 | -0.8614 | 0.0003 | -0.0792 | 0.8067 |
| ***Untergruppe 2*** | | | | | | |
| Schärfe (nach Aures) | 0.1364 | 0.6725 | 0.7594 | 0.0042 | 0.0725 | 0.8230 |
| Rauigkeit | 0.5196 | 0.0834 | -0.1161 | 0.7194 | -0.5522 | 0.0627 |
| Schwankungsstärke | 0.7260 | 0.0075 | -0.2036 | 0.5256 | -0.1800 | 0.5757 |
| Tonhaltigkeit | -0.8614 | 0.0003 | 0.1780 | 0.5800 | -0.0235 | 0.9423 |
| ***Untergruppe 3*** | | | | | | |
| $t_{Diff}$ (interaural) | -0.0093 | 0.9772 | 0.4346 | 0.1580 | 0.2806 | 0.3770 |
| $N_{DiffDir}$ (interaural) | 0.0696 | 0.8297 | 0.1590 | 0.6215 | 0.5577 | 0.0596 |
| $N_{DiffAbs}$ (interaural) | 0.4873 | 0.1081 | -0.1199 | 0.7104 | -0.3898 | 0.2103 |
| $L_{DiffDir}$ (interaural) | -0.4313 | 0.1616 | 0.1067 | 0.7413 | 0.6119 | 0.0345 |
| $L_{DiffAbs}$ (interaural) | 0.4249 | 0.1686 | -0.0803 | 0.8040 | -0.5901 | 0.0434 |
| $Corr_{Temporal}$ (interaural) | 0.2399 | 0.4526 | 0.0729 | 0.8218 | 0.6033 | 0.0378 |
| $Corr_{Spectral}$ (interaural) | 0.3488 | 0.2665 | 0.0875 | 0.7868 | 0.7853 | 0.0025 |
| Verhältnislautheit N | -0.2033 | 0.5263 | 0.7994 | 0.0018 | 0.4569 | 0.1354 |
| Schalldruckpegel $L_{SPL}$ | 0.4594 | 0.1329 | -0.5828 | 0.0468 | 0.6011 | 0.0387 |

signifikante Korrelation ($p < 0.05$)

Tabelle 5-1

94 Reuter (2005). S. 261.

95 Als Signifikanzkriterium gilt in allen Fällen ein zweiseitiges Signifikanzniveau von α = 5 %.

## 5.4 - Quantitative und qualitative Deutung der dimensionalen Struktur

Im folgenden werden die Dimensionen einzeln anhand der Ergebnisse aus den quantitativen Vergleichen mit zusätzlicher Einbeziehung von Erkenntnissen aus der qualitativen Analyse besprochen. Auf diese Weise sollen die wirksamen Einflußfaktoren ergründet werden. Da sich im MDS-Ergebnis zeigt, daß vor allem eine Trennung von Instrumenten vorliegt, wird sich der qualitative Vergleich der physikalisch-akustischen Eigenschaften maßgeblich auf Unterschiede zwischen den Instrumenten beschränken.

### 5.4.1 - Vorbetrachtung ‚unerwünschter' Einflußgrößen

Vor der eigentlichen Deutung von Dimensionen mit Hilfe der Vergleichsgrößen aus den Untergruppen 1 und 2 bietet sich an, vorher die Relevanz der Vergleichsgrößen in Untergruppe 3 zu überprüfen und zu besprechen. Nach Auslegung des Versuchsplans dürften diese Vergleichsgrößen im MDS-Ergebnis keine wichtige Rolle spielen.

Wie in Tabelle 5.1 zu sehen ist, kommt es für Vergleichsgrößen, welche dem direktionalen Hören zugeordnet werden können, nur für die interauralen Schallpegelunterschiede in dB zu statistisch signifikanten[96] Korrelationen, mit r = 0,6119 (p = 0,0345) für $L_{DiffDir}$, respektive r = -0,5901 (p = 0,0434) für $L_{DiffAbs}$. Diese Korrelationen kommen mit Dimension 3 zustande. Mit der Hinzuziehung von Streudiagrammen lässt sich die Entstehung der Korrelationen besser erklären. Hierbei ist in Abb. 5-6 und 5-7 zu sehen, daß die mittelstarken Korrelationen durch die extreme Lage des Stimulus HnD [g] begünstigt werden. Mit seiner Herausnahme verfehlen die Korrelationsanalysen das Signifikanzniveau (r = 0,4362, p = 0,1798; respektive r = -0,3975, p = 0,2261). Ferner wird aus den Streudiagrammen klar, daß die Pegeldifferenzen in den meisten Fällen deutlich weniger als 1 dB betragen, womit sie bereits unterhalb der Wahrnehmungsschwelle angesiedelt sein könnten. Es läßt sich damit für die Vergleichsgrößen des direktionalen Hörens annehmen, daß sie im Ergebnis keine Rolle spielen. Dies bestätigen auch Resultate von Kendall et al., die Unterschiede zwischen Mono- und Stereo-Wiedergabe untersuchten und keinen Unterschied für die zwei Wiedergabeformen zwischen den Bewertungen unabhängiger Gruppen finden.[97]

96 Die Verwendung des Begriffs *signifikant* bezieht sich im folgenden ausschließlich auf vorliegende statistisch signifikante Ergebnisse.

97 Kendall et al. (1998). S. 341. Es sei angemerkt, daß Kendall et al. Lautsprecher als Wiedergabequelle nutzten, während es in dieser Arbeit Kopfhörer sind.

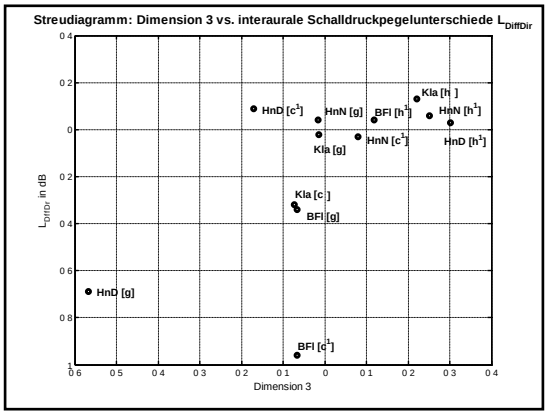


Abb. 5-6

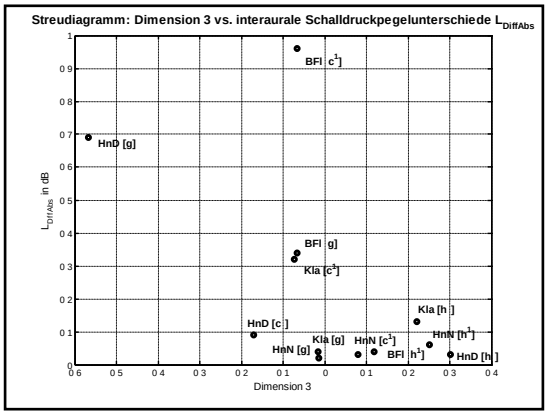


Abb. 5-7

Für die beiden verbleibenden, stereophonen Unterschieden unterliegenden Vergleichsgrößen kommt es wieder mit Dimension 3 zu signifikanten Korrelationen. Eine starke Korrelation ergibt sich mit der Vergleichsgröße der Korrelationsanalyse der interauralen Spektren ($Corr_{Spectral}$), mit r = 0,7853 (p = 0,0025); eine mittelstarke Korrelation mit dem Maximalwert der Kreuzkorrelation der interauralen Zeitsignale ($Corr_{Temporal}$), mit r = 0,6033 (p = 0,0378). Beide Korrelationen erklären sich bei Betrachtung der Streudiagramme (Abb. 5-8 und 5-9) diesmal noch eindeutiger aus den Skalenwerten des Stimulus HnD [g], der sich von allen anderen Stimuli in sowohl Dimension 3 als auch den Vergleichsgrößen am weitesten entfernt befindet. Seine Herausnahme führt dagegen zu einem deutlichen Einbruch beider Korrelationen auf nicht-signifikante Niveaus, mit r = 0,2461 (p = 0,4658), respektive r = -0,2798 (p = 0,4047). Darüber hinaus entsprechen alle Werte der interauralen Korrelation für beide Vergleichsgrößen annähernd einer perfekten Korrelation (r = 1), womit die interaurale Wahrnehmbarkeit der vorliegenden Unterschiede bereits anzuzweifeln ist. Daher wird auf eine weitere Berücksichtigung dieser Vergleichsgrößen verzichtet.

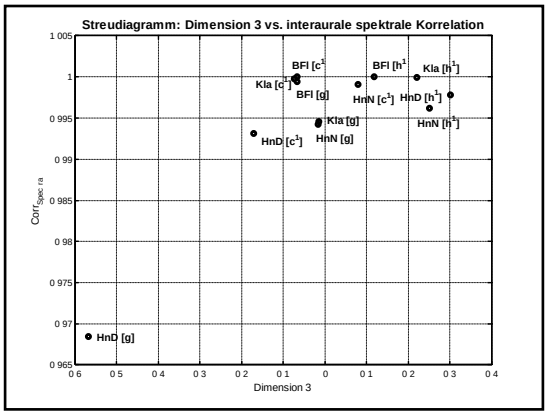


Abb. 5-8

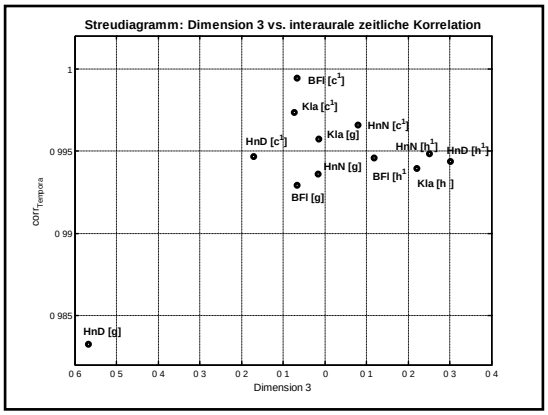


Abb. 5-9

Hinsichtlich der Verhältnislautheit *N* in Sone kommt es mit Dimension 2 zu einer starken Korrelation, mit r = 0,7994 (p = 0,0018). Dies könnte bedeuten, daß kleine Unterschiede in der Lautheit (der Umfang beträgt insgesamt 1,65 Sone) die Verteilung der Stimuli in Dimension 2 stark beeinflußt haben könnten. Auch hier empfiehlt sich, zugunsten eines differenzierten Einblicks in das Zustandekommen der Korrelation, die Betrachtung eines Streudiagramms (Abb. 5-10).

In Abschnitt 5.1 wurde bereits festgestellt, daß es in Dimension 2 zu einer deutlichen Trennung zwischen den Klarinetten-Stimuli und den übrigen Stimuli kommt. Bezogen auf die Verhältnislautheit besitzen die Klarinetten-Stimuli auch die größten Werte. Diese Werte bestimmen maßgeblich die starke Korrelation, wie sich mit der Herausnahme des lautesten Stimulus Kla [$h^1$] (r = 0,6965, p = 0,0172) bzw. aller Klarinetten-Stimuli (r = 0,0360, p = 0,9268) zeigen läßt. Weiterhin zeigt sich, daß ohne die Klarinetten-Stimuli die Lautheitswerte in keinerlei Weise Dimension 2 erklären können. Auch spiegelt sich die deutliche Abgrenzung der Klarinetten-Stimuli in Dimension 2 mit einem ‚nahtlosen' Übergang zu den weiteren Stimuli in den Lautheitswerten nur unzureichend wider. Die starke Korrelation kann zwar nicht als unbedeutend ignoriert werden, doch scheint ihr für die Verteilung der Stimuli in Dimension 2, gerade im Hinblick auf die relativ kleinen Lautheitsunterschiede, höchstens eine untergeordnete Bedeutung zuzukommen.

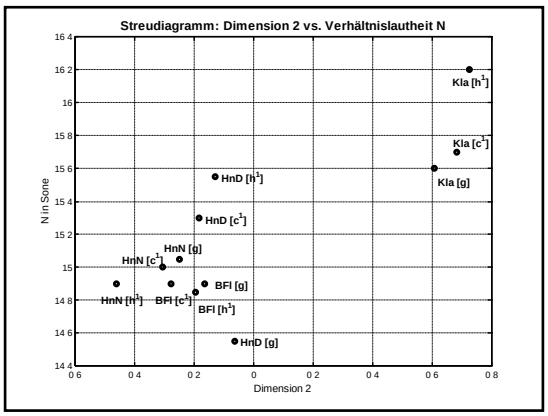


Abb. 5-10

Schwächere, doch signifikante Korrelationen ergeben sich des weiteren für den Schallpegel $L_{SPL}$ in dB. Die mittelstarke Korrelation mit Dimension 3 (r = 0,6011, p = 0,0387) erklärt sich mit Hilfe des Streudiagramms (Abb. 5-11) abermals aus der Lage des Stimulus HnD [g]. Seine Herausnahme führt auch hier zur einer nicht-signifikanten Korrelation (r = 0,2923, p = 0,3832). Mit Dimension 2 resultiert eine geringfügig schwächere Korrelation, mit r = -0,5828 (p = 0,0468). Wie im Streudiagramm (Abb. 5-12) zu sehen, ist ähnlich wie bei der Verhältnislautheit die Korrelation möglicherweise durch die Lage der Klarinetten-Stimuli bedingt. Es ist jedoch anzumerken, daß in Dimension 2 die mit der Lautstärke verbundenen Größen Verhältnislautheit und Schalldruckpegel zu konträren Korrelationen führen. Die gegensätzlichen Tendenzen beider Vergleichsgrößen können als Indiz dafür gesehen werden, daß Unterschiede in der Verhältnislautheit sowie im Schalldruckpegel keine tragende Rolle spielen.

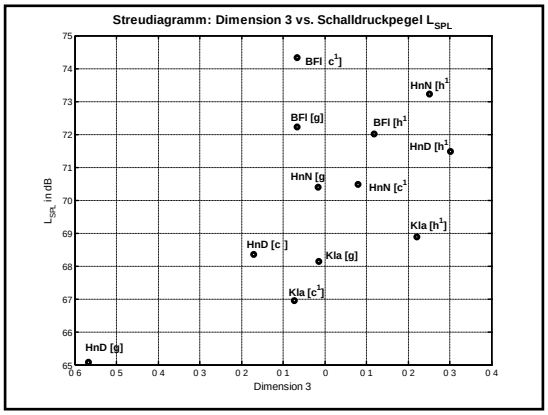


Abb. 5-11

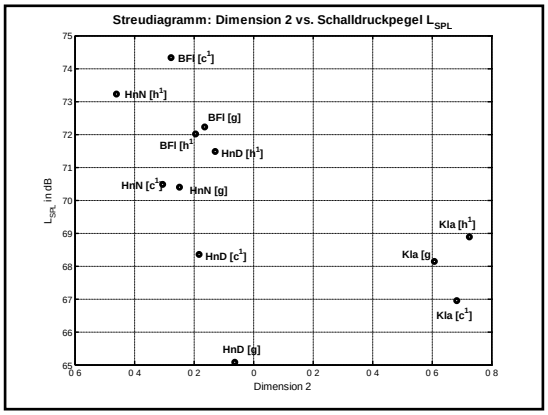


Abb. 5-12

### 5.4.2 - Diskussion und Deutung von Dimension 1

In Abschnitt 5.1 wurde festgestellt, daß in Dimension 1 eine Trennung aller Instrumente erzielt wird. Interessant ist es nun zu ergründen, welche der Vergleichsgrößen der Stimuli-Verteilung in dieser Dimension am besten entspricht. Gleichzeitig können Feststellungen aus der qualitativen Analyse die Relevanz der quantitativen Ergebnisse weiter beleuchten, sowie eventuell auch weitere, potenzielle Einflußgrößen aufzeigen.

Der Tabelle 5.1 ist zu entnehmen, daß es für Dimension 1 zu zwei signifikanten Korrelationen mit Vergleichsgrößen kommt. Die stärkste Korrelation ergibt sich mit der psychoakustischen Größe *Tonhaltigkeit*, mit r = -0,8614 (p = 0,0003) und entspricht somit einer starken negativen Korrelation. Hinsichtlich der Achsenorientierung von Dimension 1 bedeutet dies, daß angefangen bei den Horn-Stimuli die *Tonhaltigkeit* über die Klarinette, bis hin zu den Baßflöten-Stimuli stetig abnimmt. Die zweite signifikant korrelierende psychoakustische Größe, *Schwankungsstärke*, führt zu einer positiven Korrelation, mit r = 0,7260 (p = 0,0075). Für die *Schwankungsstärke* stellt sich beim Vergleich der Einzahlwerte und Zeitverläufe heraus, daß diese Korrelation sich vor allem durch zwei doppelt so hohe Werte zweier Baßflöten-Stimuli (BFl [g] und BFl [$c^1$]) ergibt. Wie zudem im Streudiagramm (Abb. 5-14) zu sehen, besitzen alle restlichen Stimuli annähernd gleich große Werte für die *Schwankungsstärke*. Des weiteren ergibt sich mit der Herausnahme der beiden oben genannten Stimuli nicht länger eine signifikante Korrelation, mit r = 0,1970 (p = 0,5855). Hierdurch läßt sich die Stimuli-Verteilung in Dimension 1 durch die *Schwankungsstärke* nur unbefriedigend erklären, da es in dieser Vergleichsgröße zu keiner Trennung unterschiedlicher Instrumente wie in Dimension 1 kommt.

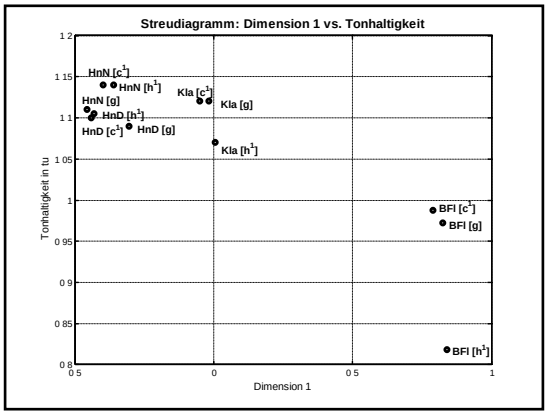


Abb. 5-13

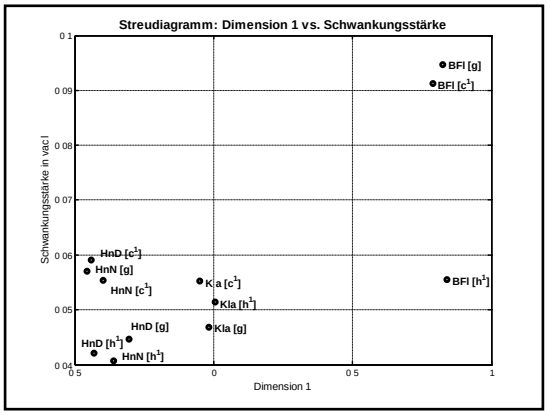


Abb. 5-14

Bezüglich der zeitlichen Verläufe der *Tonhaltigkeit* treten keine systematischen Unterschiede auf, welche die Stimuli-Darstellung in Dimension 1 zusätzlich erklären könnten, womit die Einzahlwerte letztlich als repräsentativ angesehen werden können. In Abb. 5-13 dargestelltem Streudiagramm läßt sich darüber hinaus die vorliegende Korrelation differenzierter einsehen. Hinsichtlich der Trennung nach Instrumenten läßt sich mit der *Tonhaltigkeit* die deutliche Abgrenzung der Baßflöte zur Klarinette und zum Horn gut begründen. Gleichzeitig überlappen jedoch auch die Stimuli-Werte der Klarinette und des Horns, womit sich ihre Trennung in Dimension 1 nicht vollständig begründen läßt. Es läßt sich lediglich eine Tendenz erkennen, daß die Gesamtheit der Klarinetten-Stimuli geringfügig kleinere Skalenwerte besitzt als die Gruppe aller Horn-Stimuli.

Für die Berechnung der *Tonhaltigkeit* durch das Programm *ArtemiS* liegt keine Formel vor. In der Programmhilfe-Dokumentation von *ArtemiS* findet sich jedoch eine ausführliche Beschreibung ihrer Durchführung.[98] Mittels des per FFT berechneten Spektrums werden sieben stärkste tonale Anteile (reine Töne) identifiziert, unter der Bedingung, daß alle aus unterschiedlichen Frequenzgruppen der Bark-Skala[99] stammen. Zusätzlich werden auch schmalbandige Rauschanteile kleiner einer kritischen Bandbreite identifiziert, da sie ebenfalls als tonale Anteile wirksam sind. *Tonhaltigkeit* berechnet sich danach aus dem sogenannten „Pegelüberschuss“, d.h. dem Pegel der identifizierten tonal wirksamen Komponenten, abzüglich der Ruhehörschwelle, der einzelnen Frequenzgruppenleistungen bei Nichtvorhandensein tonaler Komponenten und des breitbandigen Erregungspegels, der allein durch die tonalen Komponenten verursacht wird.

Zusammenfassend bewertet die Größe *Tonhaltigkeit* das Vorhandensein tonaler gegenüber geräuschartiger Anteile in einem Klang. Während bei der Anwendung dieser Größe zur Bewertung von maschinell erzeugten Klängen eine hohe *Tonhaltigkeit*, aufgrund einer gewissen Lästigkeit von tonalen Anteilen, eher unerwünscht ist, stellt das Vorhandensein tonaler Komponenten für

98 *ArtemiS*, Version 7. (2005). Programmhilfe: *Tonhaltigkeit*. Herzogenrath: Head Acoustics GmbH.

99 Die Bark-Skala besteht aus 24 Frequenzgruppen, welche sich von 0 bis 15500 Hz erstrecken. Die Breite einer Frequenzgruppe entspricht der sogenannten kritischen Bandbreite. Unterhalb 500 Hz besitzen diese eine konstante absolute Bandbreite von ca. 100 Hz; oberhalb 500 Hz eine konstante relative Bandbreite von 0,7 mal der Mittenfrequenz eines Bands. Die Bark-Skala entspricht einer psychoakustisch relevanten Frequenzauflösung. Sie kommt vor allem bei der Berechnung der Verhältnislautheit in Sone zum Tragen.

Instrumente geradezu eine Grundbedingung dar. Erst dadurch wird die Wahrnehmung einer diskreten Tonhöhe schließlich erst ermöglicht. Wie sich auch für die Versuchs-Stimuli zeigt, gibt es zwischen Instrumenten jedoch auch Unterschiede hinsichtlich ihrer *Tonhaltigkeit*. Die Horn-Stimuli, mit ihrer sehr breitbandigen Obertonstruktur und dem relativ hohen Amplitudenabstand zwischen tonalen und geräuschartigen Anteilen, nehmen deswegen auch die größten Werte dieser Vergleichsgröße ein. Am anderen Extrem befindet sich die Baßflöte, mit einer kaum die nötigen sieben Teiltöne aufweisenden Obertonstruktur und einem deutlich höheren Geräuschanteil. Im Vergleich mit den Horn-Stimuli erklärt sich die geringere *Tonhaltigkeit* der Klarinetten-Stimuli womöglich durch das Fehlen des 2. Teiltons, womit ein schwächer ausgeprägter Teilton in höherer Lage als siebente tonale Komponente in die Berechnung einfließt. Eine erwähnenswerte Charakteristik der Vergleichsgröße *Tonhaltigkeit* ist ihre Frequenzunabhängigkeit, da nur spektrale Amplituden und keine Frequenzen in die Berechnung einfließen. Damit eignet sie sich gerade für einen repräsentativen Vergleich von Stimuli unterschiedlicher Tonhöhe.

Bei Handel & Erickson (2004) trägt die Vergleichsgröße *spectral centroid* zu einer ansatzweisen Trennung von Instrumenten bzw. Instrumentengruppen bei. So wäre demnach im Ergebnis dieser Arbeit ein Einfluß des *spectral centroids* in Dimension 1 auch zu erwarten. Während der nicht-normierte *spectral centroid* nach Sandell ($sc_{Sandell}$) das Signifikanzniveau knapp verfehlt ($r = -0{,}5756$, $p = 0{,}0502$), führen die restlichen drei Varianten dieser Vergleichsgröße deutlich zu nicht-signifikanten Korrelationen. Trotz des knappen nicht-signifikanten Ergebnisses für $sc_{Sandell}$, wird auch hier ein Streudiagramm (Abb. 5-15) begutachtet, um zumindest das Zustandekommen der mittelstarken Korrelation zu erklären.

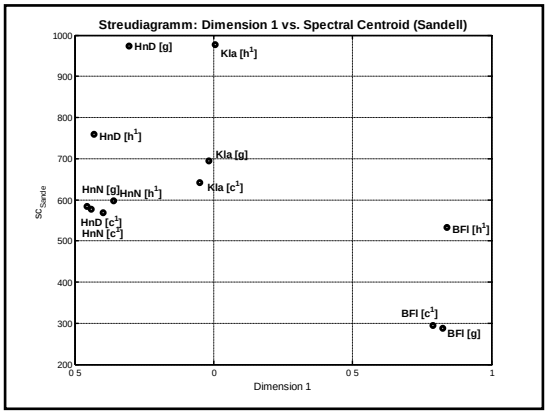


Abb. 5-15

Offensichtlich trägt nur die diskrete Trennung der Baßflöten-Stimuli von den restlichen Stimuli hinsichtlich der $sc_{Sandell}$-Werte zur Erklärung der Korrelation bei. Die Trennung zwischen Klarinetten- und Horn-Stimuli in Dimension 1 bildet sich dagegen nicht ab. Erwähnenswert ist außerdem (ausgenommen der Stimuli des Horns ohne Dämpfer) die erkenntliche Abhängigkeit nach der Tonhöhe, welche für den nicht-normierten Fall ohnehin erwartet werden konnte, da sich mit Zunahme der Grundtonfrequenz ebenfalls das Spektrum in höhere Frequenzlagen verlagert. Gerade für diese Abhängigkeit wurden die Varianten mit Normierung nach der Grundtonfrequenz miteinbezogen,

welche jedoch in beiden Fällen zu keinen signifikanten Korrelationen führen. Mit den vorliegenden Ergebnissen kann ausgeschlossen werden, daß die einzelnen Instrumente über alle Tonhöhen durch feste Verhältnisse zwischen ihrem Grundton und dem sc-Wert charakterisiert werden können, da ansonsten signifikante Korrelationen mit den normierten sc-Größen vorliegen müßten.

In den Abschnitten 2.4 und 2.5 wurden bereits die Spektral- und Temporaleigenschaften der einzelnen Instrumente besprochen, welche infolge in die qualitative Deutung miteinbezogen werden. Der deutlichste Unterschied in der Spektralcharakteristik betrifft die spektrale Ausdehnung bzw. den spektralen Reichtum der Stimuli. Für Dimension 1 läßt sich hieraus der große Abstand zwischen den Horn- sowie Klarinetten-Stimuli und den Baßflöten-Stimuli erklären, da letztere Spektren mit deutlich weniger Obertönen aufweisen. Allerdings erklärt dies wiederum nicht, weswegen beide Hornvarianten in Dimension 1 nicht zu unterscheiden sind, während die Stimuli des Horns mit Dämpfer deutlich obertonreicher sind als die ohne Dämpfer. Auch die im Vergleich zur Baßflöte geringere Trennung zwischen Horn- und Klarinetten-Stimuli läßt sich nicht eindeutig durch ihren ähnlichen spektralen Umfang begründen. Er könnte sich jedoch aus dem Fehlen des Oktavteiltons in allen Klarinetten-Stimuli erklären.

Bezüglich des Auftretens von Formanten, kann hinsichtlich einer besseren Deutung von Dimension 1 lediglich der über alle Tonhöhen vorhandene hochfrequente Formant der Baßflöte genannt werden. Die Wahrnehmbarkeit dieses Formanten kann gerade durch die schwach ausgeprägte Obertonstruktur der Baßflöten-Stimuli begünstigt werden. Dies könnte einen zusätzlichen Einfluß darauf haben, daß sich die Baßflöte so deutlich von den übrigen Instrumenten abgrenzt. Die Bedeutung dieses Formanten scheint, gerade bezogen auf den über die Tonhöhe stark variierenden Formanteigenschaften der Klarinetten-, aber auch der Horn-Stimuli, die Baßflöte von den anderen Instrumenten hervorzuheben. Ganz ähnlich läßt sich auch der Einfluß von Geräuschanteilen deuten. Auch hier kann der stark von Geräuschanteilen geprägte Baßflötenklang einen Beitrag zur großen Distanz zu den anderen Stimuli geleistet haben. Gleichzeitig wäre es nicht möglich sinnvoll zu erklären, warum die geräuschärmsten Klarinetten-Stimuli sich inmitten geräuschreicherer Stimuli wiederfinden. Daher ist ein alleiniger bzw. maßgeblicher Einfluß des Geräuschanteils für Dimension 1 nicht anzunehmen.

Abschließend würde die Berücksichtigung all dieser Charakteristiken bei der Klangfarbenbewertung eine insgesamt weit differenziertere Evaluation der Spektralcharakteristik bedeuten, als sie mit dem bloßen Vergleich mit Einzahlgrößen möglich ist. Neben der reinen spektralen Ausdehnung, welche sich noch für alle Instrumente gleichwertig mit quantitativen Methoden ausdrücken ließe, spielen auch instrumentenspezifische Eigenschaften eine Rolle: wie das Fehlen des Oktavteiltons (Klarinette), ein hoher Geräuschanteil (Baßflöte) oder auch ausgeprägte Formanteigenschaften (tiefere Tönhöhen des Horns). Im einzelnen ließen sich diese zwar auch durch quantitative Vergleichsgrößen erfassen. Da sich jedoch eine solch instrumentenspezifische Größe nicht gleichermaßen auf alle Instrumente anwenden läßt, eignen sie sich nicht für den Vergleich hinreichend unterschiedlicher Stimuli.

### 5.4.3 - Diskussion und Deutung von Dimension 2

In Dimension 2 liegt bekanntlich eine deutliche Abgrenzung zwischen den Klarinetten-Stimuli und den restlichen Stimuli vor. Der hier wirksame Einfluß vermag es weiterhin nicht zwischen Horn- und Baßflöten-Stimuli zu unterscheiden. Aus Tabelle 5.1 ist für Dimension 2 zu sehen, daß die Vergleichsgröße der Einschwingdauer ($Onset_{RMS}$) mit der dimensionalen Stimuli-Verteilung am stärksten signifikant korreliert, mit r = -0,8614 (p = 0,0003). Diese starke Korrelation ist plausibel, da schon in Abschnitt 2.5 festgestellt wurde, daß es zwischen den Klarinetten- und den restlichen Stimuli einen deutlichen Unterschied bezüglich ihrer Einschwingdauern gibt. Die negative Korrelation ergibt sich nur durch die Orientierung der Dimensionsachse, so daß mit kürzeren Einschwingdauern größere Skalenwerte in Dimension 2 einhergehen.

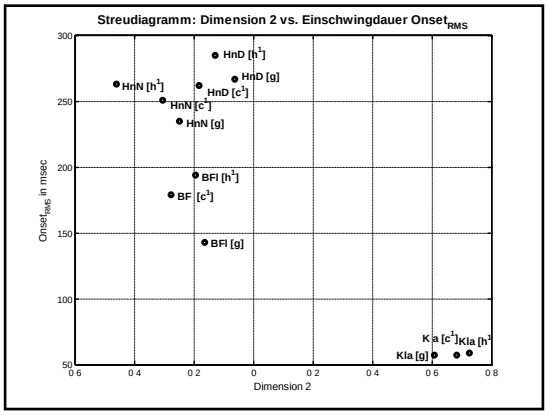


Abb. 5-16

Im Streudiagramm (Abb. 5-16) zeigt sich ferner, daß die starke Korrelation vor allem durch die deutlich kürzeren Einschwingdauern der Klarinetten-Stimuli hervorgerufen wird. Hinzu kommt, daß in den Einschwingdauern eine Trennung zwischen den Baßflöten- und allen Horn-Stimuli vorliegt, welche sich jedoch nicht auch in Dimension 2 bestätigt. Möglicherweise spielt hier die Unklarheit bei der Bestimmung der Einschwingdauern für die Horn-Stimuli eine Rolle, indem sie eventuell effektiv kürzer – d.h. den Baßflöten-Stimuli ähnlicher – sind als angenommen. Es ist allerdings auch möglich, daß bei der Wahrnehmung unterschiedlicher Einschwingdauern keine linear-skalierte Auswertung der physikalischen Dauern entscheidend ist. Ab einer hinreichend langen Einschwingdauer könnte sich beispielsweise mit der weiteren Zunahme der Dauer kein Unterschied in der subjektiven Wahrnehmung feststellen lassen. In diesem Fall würden sich Horn- und Baßflöten-Stimuli, trotz vorhandener Differenzen in ihren Einschwingdauern, in der Wahrnehmung ähnlicher sein, womit sich beide etwa gleich stark von den kürzeren Einschwingdauern der Klarinette abgrenzen.

Eine weitere Vergleichsgröße, bei der es mit Dimension 2 zu einer signifikanten Korrelation kommt, ist die psychoakustische Größe *Schärfe*, mit r = 0,7594 (p = 0,0042). Mit dieser starken positiven Korrelation könnte somit auch diese psychoakustische Vergleichsgröße als wirksame Einflußgröße in Frage kommen. Wieder wird ein Streudiagramm (Abb. 5-17) zu Rate gezogen, um zu einem besseren Verständnis über die Anpassung der Stimuli-Verteilung zu gelangen.

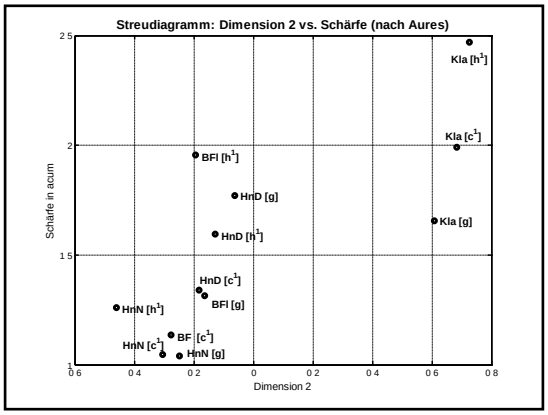


Abb. 5-17

Auch hier scheinen vergleichsweise große *Schärfe*-Werte für die Klarinetten-Stimuli die Korrelation grundlegend zu bewirken. Jedoch grenzt sich nur der höchste Klarinetten-Stimulus Kla [$h^1$] von den restlichen Stimuli ab. Die übrigen Klarinetten-Stimuli überlappen sich sogar mit einigen anderen Stimuli in einem Bereich gleicher *Schärfe*. Hierdurch läßt sich Dimension 2 nur ungenügend durch die *Schärfe* erklären, da ferner mit Herausnahme von Stimulus Kla [$h^1$] keine signifikante Korrelation mehr besteht ($r = 0{,}1792$, $p = 0{,}5981$). Bei qualitativer Evaluation der Zeitverläufe der Vergleichsgröße *Schärfe* wird des weiteren festgestellt, daß für die Klarinetten-Stimuli die zeitliche Fluktuation der *Schärfe* geringfügig kleiner ist, als bei den anderen Stimuli, womit sich Dimension 2 daraus eventuell besser erklären ließe. Zur weiteren Untersuchung wird zur quantitativen Beschreibung dieser zeitlichen Fluktuation die Standardabweichung über alle zeitlichen *Schärfe*-Werte berechnet und diese infolge einer Korrelationsanalyse mit Dimension 2 unterzogen. Diese Untersuchung liefert eine mittelstarke, negative Korrelation, welche jedoch das Signifikanzniveau verfehlt ($r = -0{,}4845$, $p = 0{,}1105$). Durch Hinzuziehung der Berechnungsformel für die *Schärfe* nach Aures[100] (siehe Formel 5.6), kann man zudem abwägen, inwiefern sich diese Vergleichsgröße überhaupt für den Vergleich von Instrumentalklängen unterschiedlicher Tonhöhe eignet.

$$S = K_a \cdot \frac{\int_0^{24\,Bark} n'(z) \cdot e^{0{,}171 \cdot z}\, dz}{\log\left(\frac{N}{20} + 1\right)} \qquad \text{Formel 5.6}$$

*S: Schärfe nach Aures; $K_a$: Normierungsfaktor; n'(z): spezifische Lautheit pro Frequenzgruppe; z: Frequenzgruppe in Bark; N: Gesamtlautheit*

Im Nenner befindet sich ein nicht-frequenzabhängiger Term, welcher zur Normierung auf die Gesamtlautheit *N* dient. Dagegen ist der Ausdruck im Zähler deutlich frequenzabhängig, da mit zunehmender Frequenz bzw. Frequenzgruppe *z* eine exponentiell stärkere Gewichtung höherfrequenter Anteile erfolgt. Somit nimmt *Schärfe* bei Vorhandensein stark hochfrequenter Anteile zu. Bei einem Vergleich von Instrumentalklängen gleicher Tonhöhe könnte ein Unterschied in der *Schärfe*

100 *ArtemiS*, Version 7. (2005). Programmhilfe: *Schärfe*. Herzogenrath: Head Acoustics GmbH.

für die Unterscheidbarkeit bzw. Ähnlichkeit von Klängen eine Rolle spielen. Bei Vorhandensein unterschiedlicher Tonhöhen resultieren jedoch, wie in dieser Arbeit, für ein und dasselbe Instrument aufgrund des Tonhöhenunterschieds unterschiedliche *Schärfe*-Werte. Deswegen erscheint eine Trennung von Instrumenten anhand dieser Größe nicht realisierbar. Wie im Streudiagramm (Abb. 5-17) ferner zu sehen ist, spielt für die *Schärfe*-Werte eines Instruments offenbar nicht nur die Tonhöhe eine Rolle. Denn für die Baßflöte gibt es z.B. keine monoton steigende Beziehung zwischen Tonhöhe und *Schärfe*. Schließlich kann bereits aus der Berechnungsmethode angenommen werden, daß sich die *Schärfe* zur Deutung eines MDS-Ergebnis von Instrumentalklängen unterschiedlicher Tonhöhe weniger eignet. Dies wird auch durch die keinesfalls zur Erklärung von Dimension 2 beitragende Korrelation (ohne Stimulus Kla [$h^1$]) quantitativ bestätigt.

Mit der Einbeziehung qualitativer Erkenntnisse aus den Abschnitten 2.4 und 2.5 zur Deutung von Dimension 2 können einige Spektraleigenschaften mit der Trennung der Klarinette von den anderen Instrumenten in Verbindung gebracht werden. Keine dieser Spektraleigenschaften läßt sich jedoch über alle Tonhöhen gleich gut zur Erklärung von Dimension 2 heranziehen.

Aus den Formanteigenschaften könnte man schließen, daß die stärker von Formanten geprägten Klänge der Baßflöte und des Horns sich von der Klarinette, die eine über die Tonhöhen inkonsistente Prägung durch Formanten aufweist, abgrenzen. Für das Horn läßt die Formantprägung jedoch mit zunehmender Tonhöhe stetig nach, womit sich die höheren Horn-Stimuli wiederum denen der Klarinette in ihrer Ähnlichkeit annähern müßten, was sich in Dimension 2 nicht bestätigt. Ferner könnte die einzigartige Spektralcharakteristik der Klarinette zu dieser Trennung beitragen. Da mit zunehmender Tonhöhe die schwache Ausprägung der geradzahligen Teiltöne oberhalb des dritten Teiltons bereits stetig nachläßt, verbleibt als gemeinsames Charakteristikum aller Klarinetten-Stimuli lediglich die schwache Ausprägung des zweiten Teiltons (Oktavton). Es läßt sich folgern, daß sowohl die Formanteigenschaft der Horn- und Baßflöten-Stimuli, als auch die besondere Spektralcharakteristik der Klarinette aufgrund ihrer Frequenzabhängigkeit nicht als Haupteinflußquellen für die Stimuli-Verteilung in Dimension 2 in Frage kommen. Außerdem besitzen die Klarinetten-Stimuli den geringsten Geräuschanteil, womit sich ihre deutliche Abgrenzung von der Baßflöte erklären ließe. Dieser Logik nach müßten sich aber gleichzeitig die Horn-Stimuli (mit mittlerem Geräuschanteil) in Dimension 2 zwischen diesen beiden Extremen befinden und nicht auf gleicher Höhe mit denen der Baßflöte. Durch eine qualitative Betrachtung der Spektraleigenschaften läßt sich somit keine bessere bzw. zufriedenstellende Deutung von Dimension 2 erreichen.

Schließlich deutet einiges darauf hin, daß sich die genannten Spektraleigenschaften gegenüber den temporalen Eigenschaften weitaus weniger eignen, Dimension 2 zu erklären. Wie sich auch für die quantitative Vergleichsgröße der Einschwingdauer aus dem Effektivwertverlauf zeigt, wird über die unterschiedlichen Einschwingdauern die Trennung der Klarinette sehr gut begründet. Jedoch bleibt der Einwand, daß sich Horn- und Baßflöten-Stimuli trotz unterschiedlicher Einschwingdauern in Dimension 2 praktisch nicht unterscheiden. Stattdessen kann angenommen werden, daß die Stimuli-Verteilung in Dimension 2 möglicherweise durch eine weniger differenzierte Auswertung der Dauer

erzielt wurde: im einfachsten Fall durch die Wahl zwischen zwei ordinalen Kategorien, für *lange* und *kurze* Dauern.

Weiterhin beachtenswert ist der im subjektiven Hörvergleich der Stimuli wahrnehmbare Unterschied in der stereophonen Wiedergabe. Die größten Unterschiede gibt es hier vor allem zwischen den Klarinetten- und den anderen Stimuli. Hierbei werden die anderen Stimuli insgesamt ‚räumlicher' bzw. ‚breiter' in der binauralen Wiedergabe abgebildet. Horn- und Baßflöten-Stimuli besitzen diesbezüglich einen sehr ähnlichen Klangeindruck, was der Stimuli-Verteilung in Dimension 2 gut entsprechen würde. Mit einigen Vergleichsgrößen der Untergruppe 3 wurde versucht, diesen Unterschied zu erklären. Wie sich in Abschnitt 4.5.1 gezeigt hat, lassen sich hierfür weder die Größen aus dem direktionalen Hören, noch die Korrelationsanalysen der Zeitsignale und Spektren für beide Hörkanäle als Erklärung nutzen. Stattdessen korrelieren diese Größen bekanntlich mit Dimension 3. So bleibt anzunehmen, daß dieser auditive Unterschied zwischen der Klarinette und den anderen Instrumenten auf noch komplexeren Unterschieden in den Zeitsignalen bzw. Spektren basiert, als sie mit den durchgeführten Vergleichsgrößen erfaßt werden können. Leider kann diese bleibende Unklarheit im Rahmen dieser Arbeit nicht mehr weiter untersucht werden, weil dies zu umfangreich wäre.

In Ergänzung weisen Klänge von Klarinetten gegenüber anderen Instrumenten oftmals eine subjektive ‚Nähe' bzw. ‚Direktheit' auf; ganz unabhängig von ihrer eigentlichen räumlichen Position, sowie auch raumakustischen Einflüssen. Dieses deutet letztlich jedoch auf die oben genannten Unterschiede, welche zwar subjektiv registriert, jedoch nicht physikalisch-akustisch identifiziert werden können. In einem Vergleich des oberen Registers zwischen Klarinette und großer C-Flöte äußert sich Meyer, daß die Flöte „in dieser Lage zwar einen ähnlichen Obertonaufbau zeigt, aber erheblich höhere Rauschanteile liefert und zudem stärkere Fluktuationen in der zeitlichen Mikrostrukur des Pegelverlaufes aufweist, so daß der Klang der Klarinette ruhiger und fester und dadurch auch kräftiger wirkt"[101]. Diese Erkenntnis liefert möglicherweise einen Anhaltspunkt dafür, wo sich diese ‚Nähe' bzw. ‚Direktheit' der Klarinette physikalisch nachweisen lassen könnte. Zusätzlich läßt sich die Hypothese aufstellen, daß es sich bei der vorliegenden Trennung auch um eine kategoriale Trennung nach dem Material der Instrumente handeln könnte: Metall (für Baßflöte und Horn) und Holz (für Klarinette). Dies beschränkt sich nicht nur auf semantische Assoziationen bezüglich der Instrumente, sondern auch auf akustisch wahrnehmbare Unterschiede, welche durch diese grundlegend unterschiedlichen Materiale hervorgerufen werden könnten.

### 5.4.4 - Diskussion und Deutung von Dimension 3

Für Dimension 3 ergibt sich die stärkste signifikante Korrelation aus den Vergleichsgrößen des Tonhöhenunterschieds, mit $r \approx 0{,}74$ ($p < 0{,}006$). Hierbei führen sowohl Tonhöhenunterschiede in Cents als auch die Frequenzverhältnisse $f/f_R$ zu annähernd gleich starken Korrelationen. Auf dem ersten Blick überrascht diese Feststellung, da es mathematisch einen funktionalen Unterschied zwischen der reinen Frequenzverhältnis-Skala ($f/f_R$) und der logarithmierten Skala in Cents ($\Delta p$) gibt,

101 Meyer (1999). S. 66.

woraus auch grundlegend unterschiedliche Korrelationen hervorgehen müßten. Es stellt sich jedoch bei genauerer Untersuchung heraus, daß sich im vorhandenen Wertebereich der Frequenzverhältnisse ($f/f_R$ = {1,0 ... 2,5}) die beiden Funktionsverläufe annähernd linear zueinander verhalten. Es bestätigt sich also auch durch objektive Kriterien gewissermaßen ein Einfluß der Tonhöhe für Dimension 3. Durch die annähernd gleichen Korrelationen läßt sich jedoch nicht mit Sicherheit auf die Cents-Skala oder den Frequenzverhältnissen als Skalierungskorrelat für den Tonhöheneinfluß schließen. So bleibt unklar, ob die Skalierung in Dimension 3 auf musikalisch relevante Intervallgrößen oder auf physikalische Verhältnisse der Grundtonfrequenzen zurückzuführen ist.

Die Korrelation mit den Vergleichsgrößen der Tonhöhenverhältnisse fallen insgesamt nicht so stark aus, wie die stärksten Korrelationen in den anderen Dimensionen. Dies ist bezüglich der dimensionalen Abbildung auch nachzuvollziehen, da die bekannten Inkonsistenzen in der Abbildung der Tonhöhenbeziehung g–$c^1$ zur Schwächung der Korrelationen beitragen. Dieses läßt sich in einem Streudiagramm für die Tonhöhenunterschiede in Cents (Abb. 5-18) einsehen. Für Baßflöte und Klarinette kommt es bezüglich der Tonhöhen g und $c^1$ zu einem nicht-linearen intra-instrumentalen Verlauf, während sich für beide Instrumentalvariationen des Horns ein linearer Verlauf über alle Tonhöhen klar erkennen läßt. Überdies liegen, wie in Abschnitt 5.1 bereits erwähnt, unterschiedliche Gewichtungen bzw. Skalierungen des Tonhöheneinflusses innerhalb der Instrumente bzw. Instrumentalvariationen vor. Es sei erwähnt, daß dies im MDS-Ergebnis von Handel & Erickson (2004) ebenfalls der Fall ist und somit nicht als untypisch angesehen werden muß.

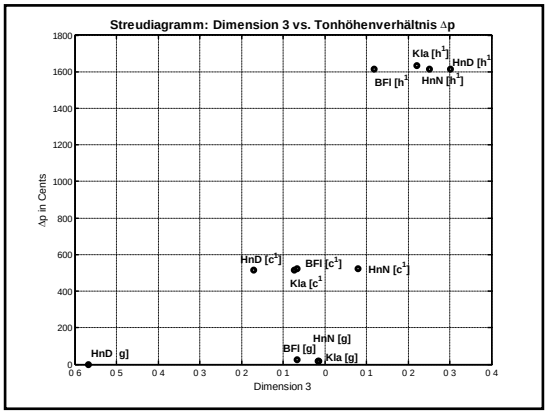


Abb. 5-18

Die genannten Abweichungen von einer idealen, linearen Abbildung des Tonhöheneinflusses lassen offen, wodurch sie genau bedingt sind. Anzunehmen ist die bereits genannte Möglichkeit der interindividuellen Bewertungsvariabilität oder daß im Vergleich zur höchsten Tonhöhe $h^1$ das relativ kleine Intervall der Quarte g–$c^1$ weniger bzw. sogar unbeeinflußt von Tonhöhenunterschieden bleibt. Auch darf nicht ausgeschlossen werden, daß die Tonhöhen bzw. -lagen im Hinblick auf ihr instrumentenspezifisch unübliches Auftreten im Musikrepertoire einen zusätzlichen Faktor darstellen. Die Baßflöte als recht selten gehörtes Instrument bzw. das ebenfalls selten genutzte unterste Register der Klarinette könnten hier beispielsweise mit dem Tonhöheneinfluß negativ interagiert haben. In Anbetracht der instrumentenabhängigen Ausbreitung der Stimuli in

Dimension 3 ist es möglich, daß neben einem effektiven Tonhöheneinfluß ein weiterer Einfluß, in Form einer positiven bzw. unterstützenden Interaktion, vorzufinden ist. Der klare Gegensatz in der Ausbreitung der zwei Hornvarianten in Dimension 3 kann damit in Verbindung gebracht werden, indem zwischen den Varianten bestehende klangliche Unterschiede darin abgebildet werden könnten. Wie in Abschnitt 2.5.1 dargestellt, existieren bezüglich der spektralen Ausdehnung deutliche Unterschiede zwischen den Hornvarianten. Da es innerhalb von Dimensionen 1 und 2 lediglich zu einer Trennung der drei Instrumente kommt, ist es möglich, daß diese Unterschiede zusätzlich Dimension 3 beeinflussen.

Weitere eventuell relevante Korrelationen zwischen Vergleichsgrößen und Dimension 3 treten unerwarteterweise mit einigen Vergleichsgrößen der Spektraleigenschaften auf. So ergibt sich für beide Varianten des normierten *spectral centroid* eine annähernd gleichstarke, signifikante Korrelation wie mit den Tonhöhengrößen, mit $r \approx -0{,}72$ ($p < 0{,}009$). Im Falle von $sc_{Sandell\text{-}f0}$ stellt sich bei Betrachtung des Streudiagramms (Abb. 5-19) heraus, daß der Stimulus HnD [g] diese Korrelation maßgeblich begünstigt. Mit seiner Herausnahme verringert sich die Stärke der Korrelation auf ein nicht-signifikantes Niveau, mit $r = -0{,}3582$ ($p = 0{,}2795$). Dieses gilt jedoch nicht für $sc_{FFT\text{-}f0}$ (Abb. 5-20), dessen Korrelation mit der Herausnahme von Stimulus HnD [g] sich sogar leicht verstärkt. Beim Verfolgen des Verlaufs der einzelnen Tonhöhen pro Instrument bzw. Instrumentalvariation ist zu erkennen, daß die dazugehörigen normierten sc-Werte mit Zunahme der Tonhöhe stetig abnehmen, d.h. das spektrale Zentrum nähert sich dem Grundton. Dies ist für alle Instrumente durchgängig festzustellen und bei $sc_{Sandell\text{-}f0}$ sehr ausgeprägt, bei $sc_{FFT\text{-}f0}$ ansatzweise zu erkennen. Es zeichnet sich letztlich in der Vergleichsgröße des normierten *spectral centroids* ebenfalls ein konsistenter Tonhöheneinfluß ab. Daraus erklärt sich wiederum die hohe Korrelation mit Dimension 3.

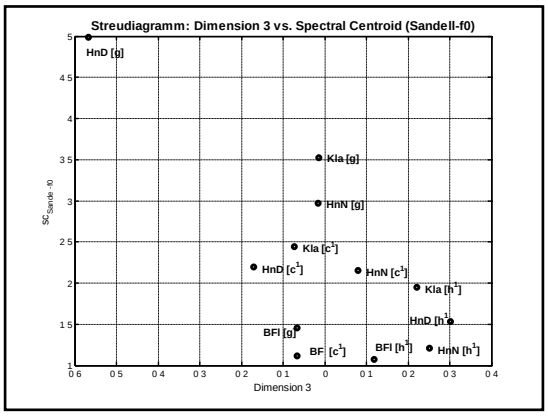


Abb. 5-19

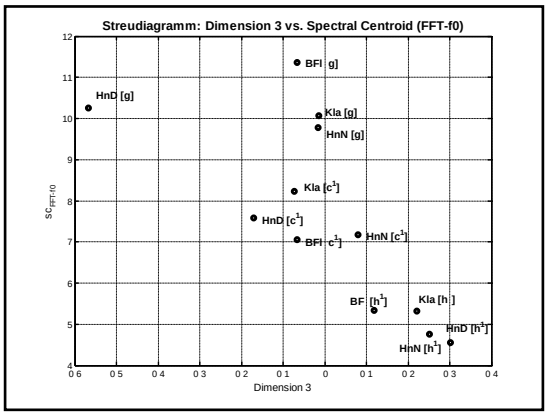


Abb. 5-20

Neben *spectral centroid* kommt es für die Vergleichsgröße *spectral variability* (sv) ebenfalls zu signifikanten Korrelationen mit Dimension 3, wenngleich schwächer. Diese Korrelationen sind für beide Zeitauflösungen, wie auch für beide Spektraldarstellungen, annähernd gleich stark. Für $sv_{Sandell}$ ergibt sich eine mittelstarke negative Korrelation, mit $r = -0{,}6508$ ($p = 0{,}0219$), während sich für $sv_{FFT}$ eine gleichstarke, jedoch positive Korrelation ergibt, mit $r = 0{,}6202$ ($p = 0{,}0314$).[102]

102 Aus praktischen Gründen wird an dieser Stelle nur die Zeitauflösung von 85 msec berücksichtigt.

Diese gegensätzliche Orientierung der Korrelationen, sowie die deutlich unterschiedliche Größe der Werte, lassen sich dadurch erklären, daß $sc_{Sandell}$-Werte sich im Bereich einiger Hundert Hz bewegen, während $sc_{FFT}$-Werte im vierstelligen Hz-Bereich angesiedelt sind und sich beide Größen somit auf zwei unterschiedliche Frequenzbereiche beziehen. Diese Unterschiede ergeben sich aus den grundlegend unterschiedlichen Frequenzauflösungen der beiden Spektralanalysen, denen die sv-Varianten zugrunde liegen.

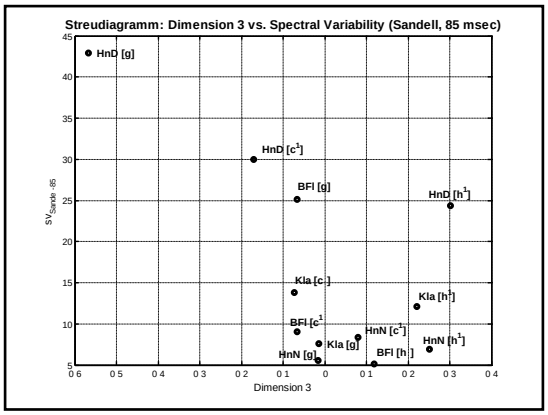


Abb. 5-21

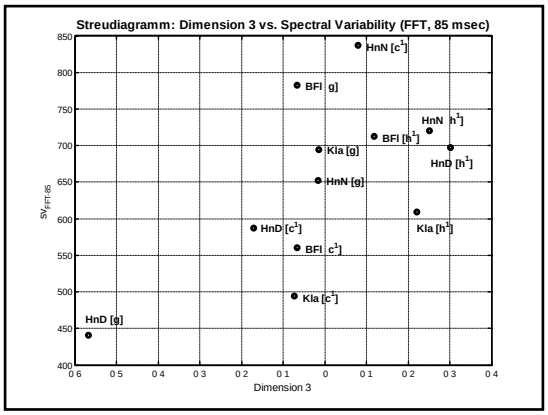


Abb. 5-22

Die gegensätzliche Orientierung bedeutet auch, daß es zwischen dem niederfrequenten und den höherfrequenten Bereichen gegensätzliche Tendenzen bezüglich der spektralen Fluktuation gibt. Besonders erwähnenswert ist, daß der Stimulus HnD [g] auch hier stark zur vorhandenen Korrelation beiträgt, ganz egal welcher Orientierung (Abb. 5-21 und 5-22). Für die tieferfrequente $sv_{Sandell}$ besitzt er mit Abstand die höchste spektrale Fluktuation, während er in den höherfrequent angesiedelten $sv_{FFT}$-Werten die geringste Fluktuation besitzt. Mit seiner Herausnahme verfehlen beide Korrelationen das Signifikanzniveau, mit r = -0,2304 (p = 0,4954), respektive r = 0,3487 (p = 0,2932). Demzufolge nimmt Stimulus HnD [g] eine einzigartige Stellung ein, in dem er gegensätzliche Extrema in beiden sv-Varianten bildet. Für die Erklärung von Dimension 3 lassen sich diese Vergleichsgrößen zwar nicht umfassend nutzen. Jedoch läßt sich möglicherweise mit der exponierten Stellung des Stimulus HnD [g] dessen extreme Position in Dimension 3 erklären, was sich wie bekannt nicht mit den Vergleichsgrößen des Tonhöhenunterschieds bewerkstelligen läßt. Bemerkenswert ist auch, daß derselbe Stimulus HnD [g] schon bei einigen der Vergleichsgrößen der Untergruppe 3 eine Sonderstellung innehat. Mit Rückblick auf die qualitative Auswertung der physikalisch-akustischen Eigenschaften (in den Abschnitten 2.4 und 2.5) läßt sich diese Sonderstellung des Stimulus HnD [g] weiter erklären. Im Vergleich zu den Spektren der übrigen Horn-Stimuli besitzt er geringfügig schwächere Amplitudenwerte. Außerdem stellt er einen von zwei Stimuli dar, bei denen die Existenz von Vorläuferimpulsen im Blasansatz nachgewiesen werden konnte. Diesbezüglich besitzt er aufgrund des obertonreicheren Spektrums des Horns mit Dämpfer auch stärker ausgeprägte breitbandige Anteile im Blasansatz (als der Stimulus HnN [g]). Somit lassen sich auch hierin Anhaltspunkte für seine Sonderstellung in Dimension 3 finden.

Angesichts der in Abschnitt 4.4.3 genannten Einschränkung bezüglich der Interpretierbarkeit des

MDS-Ergebnisses für Dimension 3 müssen die oben gestellten Annahmen als empirisch nicht abgesichert eingestuft werden. Als wahrscheinlich darf erstens ein allgemeiner Tonhöheneinfluß, ohne Rückschluß auf eine skalierbare Größe, sowie zweitens die Möglichkeit einer Sonderstellung des Stimulus HnD [g], angenommen werden.

### 5.4.5 - Zusammenfassung

Die Trennung aller Instrumente in Dimension 1 kann mit einiger Sicherheit auf spektrale Eigenschaften zurückgeführt werden. Mit der starken Korrelation der Vergleichsgröße *Tonhaltigkeit* kann angenommen werden, daß sich der Einfluß anteilig auf die frequenzmäßige Ausdehnung bzw. Anzahl von Teiltönen sowie des relativ dazu vorhandenen Geräuschanteils stützt. Dieses läßt sich auch problemlos mit den Erkenntnissen aus der qualitativen Analyse der Spektren in Einklang bringen. Zudem existiert die Möglichkeit, daß die Ähnlichkeitsbewertungen der VP auf mehreren, nicht quantitativ zusammenfassbaren Eigenschaften beruhten. So läßt sich die deutlichere Trennung der Baßflöte von den anderen Instrumenten vielleicht auf eine Anzahl charakteristischer spektraler Unterschiede (erhöhter Geräuschanteil, begrenztes Obertonspektrum, sich darüber befindlicher Formant) zurückführen. Eine so hohe Vielfalt an Charakteristiken läßt sich womöglich gar nicht erst durch eine skalierbare Einzahlwert-Größe ausdrücken, woraus auch keine direkte Berechnung von Korrelationen möglich wäre.

Die Trennung der Klarinetten-Stimuli von denen des Horns und der Baßflöte in Dimension 2 ist mehrdeutig zu interpretieren. Am besten läßt sich diese Trennung mit einer starken Korrelation auf die unterschiedlichen Einschwingdauern zurückführen. Auch hier scheint entweder ein weiterer Einfluß eine Rolle zu spielen, indem vorhandene Unterschiede in den Einschwingdauern von Horn- und Baßflöten-Stimuli in der dimensionalen Abbildung angeglichen werden oder diese Angleichung der Einschwingdauern liegt in einer einfachen ordinalen Kategorisierung der Einschwingdauern (in *lang* und *kurz*) begründet. Alternativ könnte auch der in der binauralen Wiedergabe wahrnehmbare subjektive Unterschied im Räumlichkeitseindruck zwischen Klarinetten- und den übrigen Stimuli einen Einfluß haben. Dies läßt sich jedoch nicht mittels signifikant korrelierender Vergleichsgrößen quantitativ nachweisen. Gerade im Hinblick auf die besprochene empfundene ‚Nähe' bzw. ‚Direktheit' von Klarinettenklängen sollte dies aber nicht vernachlässigt werden. Daran angeknüpft könnte auch eine Kategorisierung entsprechend des Materials der Instrumente (Metall bzw. Holz) einen Einfluß haben. Letztlich würde eine positive Interaktion von beiden potenziellen Einflußgrößen auch in Frage kommen, was sich jedoch im Rahmen dieser Arbeit nicht weiter überprüfen läßt.

Mit einiger Sicherheit kann für Dimension 3 ein wirksamer Einfluß der Tonhöhe angenommen werden. Jedoch beschränkt sich diese Annahme auf einen relativ grob definierten Einfluß, ohne eine Aussage zu machen inwiefern er skalierbar ist bzw. welche der verwendeten Vergleichsgrößen letztlich den Ausschlag gibt. Dies ist einerseits durch die Unstimmigkeiten in der dimensionalen Abbildung der Tonhöhen g und $c^1$ für Baßflöte und Klarinette, andererseits jedoch vor allem durch die in Abschnitt 4.4.3 genannte eingeschränkte und wenig abgesicherte Aussagekraft für Dimension 3 bedingt. Zusätzlich läßt sich die exponierte Position des Stimulus HnD [g] in der geometrischen

Konstellation X nicht auf einen Einflußbereich eingrenzen. Denn wie sich zeigt, besitzt dieser Stimulus sowohl in der Größe *spectral variability*, als auch in mehreren der Stereophonie unterliegenden Vergleichsgrößen eine prägnante Sonderstellung. In den Erkenntnissen aus den physikalisch-akustischen Analysen lassen sich zudem Anhaltspunkte für diese Sonderstellung finden. Gerade für Dimension 3 wäre für eine Absicherung der Annahmen, sowie informativerer Ergebnisse, eine Reproduktion des Versuchs mit Hinzunahme weiterer Tonhöhen empfehlenswert.

In Abschnitt 5.2 wurde für die Deutung von Dimensionen ein Ansatz gewählt, der sich auf sowohl quantitative als auch qualitative Methoden stützt. In diesem Kapitel zeigt sich, daß sich dieser Ansatz bewährt hat. Aus quantitativen Vergleichen ergeben sich viele signifikante Korrelationen mit Vergleichsgrößen. Für die meisten Fälle stellt sich bei der qualitativen Auswertung von Streudiagrammen allerdings heraus, daß sie sich nur ungenügend zur Erklärung der dimensionalen Struktur eignen. Diese Häufung von irrelevanten Korrelationen ergibt sich aus der relativ geringen Anzahl von Stimuli. Bei der Anzahl von 12 Stimuli läßt sich allein durch die extreme Position eines einzelnen Stimulus oftmals eine starke signifikante Korrelation erhalten. Daher ist eine Evaluation von Streudiagrammen in dieser Situation unumgänglich, um nicht fälschlicherweise irrelevante Korrelationen zur Deutung einer Dimension zu verwenden. Außerdem zeigt sich, daß Erkenntnisse aus der qualitativen Auswertung der physikalisch-akustischen Analysen zur Diskussion und Deutung von Dimensionen beitragen können. Auch nicht-skalierbare oder instrumentenspezifische Eigenschaften können sich dadurch identifizieren lassen. Ferner wird die Betrachtung auf die Berücksichtigung von kategorialen Unterschieden ausgeweitet.

# 6 - DISKUSSION & VERGLEICH DES MDS-ERGEBNISSES

Nachdem das MDS-Ergebnis ausführlich ausgewertet und in seiner dimensionalen Bedeutung interpretiert wurde, werden infolge die Ergebnisse dieser Arbeit mit denen von Handel & Erickson (2004) sowie auch anderen relevanten Forschungsergebnissen verglichen und auf ihre Bedeutung hin diskutiert.

Eine direkte Vergleichbarkeit bieten nur die Versuchsergebnisse von Handel & Erickson (2004), da der Versuch dieser Arbeit an ihrem Versuchsmodell orientiert ist. Gründe für eine eingeschränkte Vergleichbarkeit zu anderen MDS-Untersuchungen ergeben sich aus der Tatsache, daß diese Versuche meist nur mit einer festen Tonhöhe durchgeführt wurden, woraus ein Tonhöheneinfluß gar nicht erst möglich ist. Als Resultat ergeben sich im dimensionalen Ergebnis somit grundlegend verschiedene Organisationsstukturen. Allerdings werden einige MDS-Untersuchungen berücksichtigt, welche den Tonhöheneinfluß gezielt mit in ihren Versuchsplan einbeziehen. Diese weisen zwar keine direkte Vergleichbarkeit zum Versuch dieser Arbeit auf, jedoch können sie bei der Diskussion des Tonhöheneinflusses einbezogen werden.

## 6.1 - Vergleich des vorliegenden Ergebnisses mit Handel & Erickson (2004)

Wie eingangs in Abschnitt 1.2.1 bereits vorgestellt, liegt in der dreidimensionalen MDS-Konstellation von Handel & Erickson (2004) strukturgebend ein Einfluß der Tonhöhe vor. Dieser Einfluß läßt sich in Dimension 1 klar nachweisen. Im MDS-Ergebnis dieser Arbeit läßt sich ein Tonhöheneinfluß erst mit Hinzunahme von Dimension 3 erkennen. Zudem ist dieser Einfluß weitaus weniger ausgeprägt und besitzt ferner eine geringere Bedeutung für die Gesamtstruktur der geometrischen Konstellation X.

Im Ergebnis von Handel & Erickson (2004) liegt weiterhin eine Tendenz zur Trennung von Instrumenten in Dimension 2 vor. Doch kommt es im dimensionalen Verlauf auch zu Überschneidungen der einzelnen Instrumente, womit in ihrem Ergebnis keine diskrete Trennung aller Instrumente erreicht wird. Im Gegensatz dazu liefert das MDS-Ergebnis dieser Arbeit eine eindeutig diskrete räumliche Trennung der Instrumente in einzelne Stimuli-Gruppen. Diese Trennung basiert zudem auf einer Differenzierung in zwei Dimensionen, wobei in Dimension 2 nur eine Trennung der Klarinette von den anderen Instrumenten erzielt wird. Während sich bei Handel & Erickson (2004) die Vergleichsgröße *spectral centroid* als korrelierende Größe für die ansatzweise Trennung von Instrumenten annehmen läßt, trägt diese Vergleichsgröße (in insgesamt vier Berechnungsvarianten) im MDS-Ergebnis dieser Arbeit nicht zur Instrumententrennung bei. Dagegen beweist sich der bei Handel & Erickson (2004) als weniger bedeutsam eingestufte Einfluß der Einschwingdauer als signifikant korrelierende Größe für Dimension 2. Die Trennung aller Instrumente in Dimension 1 läßt sich darüber hinaus am ehesten mit der von Handel & Erickson (2004) nicht verwendeten psychoakustischen Größe *Tonhaltigkeit* in Einklang bringen. Dimension 3 führt bei Handel & Erickson (2004) zu einer nur mittelstarken Korrelation mit einer Vergleichsgröße des ‚Vibrato-Umfangs'. Die Autoren griffen bei ihren Stimuli auf bereits aufgenommene Instrumentalklänge zurück, in welchem Fall das Vorhandensein von Vibrato

nicht von vornherein ausgeschlossen werden konnte. In dieser Arbeit war bei der Aufzeichnung der Stimuli jedoch ausdrücklich kein Vibrato erwünscht, womit auf eine nachträgliche Untersuchung eines eventuellen Einflusses verzichtet wird.

Zusammenfassend zeigt sich in beiden Versuchen ein Einfluß der Tonhöhe, wenn auch von grundlegend unterschiedlicher Bedeutung für die MDS-Struktur. Des weiteren basiert die Trennung nach Instrumenten in beiden Fällen auf spektralen Eigenschaften der Stimuli. In dieser Arbeit kann zudem die Einschwingdauer für die (mehrdimensionale) Trennung einbezogen werden.

## 6.2 - Diskussion der Unterschiede zwischen den Versuchen

In MDS-Versuchen mit einer Tonhöhe stellt bereits Grey fest, daß sowohl spektrale als auch temporale Eigenschaften der Klänge zur dimensionalen MDS-Abbildung beitragen.[103] Aus dem Vorhandensein dieser beiden Einflußquellen in seinen Ergebnissen folgert Lakatos, daß auf der Basisebene die menschliche Repräsentation der Klangfarbenwahrnehmung bidimensional ist.[104] Diese hypothetische Basisebene ist letztlich aber stark vom Gesamtkontext der eingesetzten Stimuli abhängig. So können sich bereits spektrale Eigenschaften in mehreren Dimensionen abbilden, wie beispielsweise die Vergleichsgrößen *spectral centroid* und *spectral variability*. Der Einfluß temporaler Eigenschaften ist von der Stimuli-Zusammensetzung abhängig. Während bei gleichzeitigem Vorkommen von perkussiven und kontinuierlich angeregten Klängen ein Einfluß sehr wahrscheinlich ist, muß dies bei einer starken Ähnlichkeit temporaler Charakteristiken von Stimuli wiederum nicht zutreffen. Des weiteren folgern Melara et al. aus ihrem Versuch mit synthetischen Stimuli auf einen ausschließlichen Einfluß von spektralen Anteilen für die Klangfarbenwahrnehmung[105], während Iverson in einem Versuch zur Untersuchung der Rolle der Klangfarbe in der *Auditory Stream Segregation* feststellt, daß temporale Eigenschaften der Klangfarbe maßgeblich zur Trennung von ‚Hörströmen' beitragen[106]. Die dimensionale Abbildung einer MDS-Untersuchung zur Klangfarbe ist demnach stark vom Kontext der jeweilig vorliegenden Stimuli-Zusammensetzung sowie -Darbietung abhängig.

Zwischen den Ergebnissen dieser Arbeit und denen von Handel & Erickson (2004) gibt es, wie bereits dargestellt, den grundlegenden Unterschied bezüglich der Trennung der Instrumente, indem es in dieser Arbeit zu einer klaren räumlichen Trennung aller Instrumente kommt. Als wesentliche Frage bleibt zu klären, wie sich die Tatsache erklären und begründen läßt, daß bei Handel & Erickson (2004) bei einem vergleichbaren Versuchsplan keine derartig klare Trennung erreicht wird. Die Antwort läßt sich am plausibelsten in den Unterschieden zwischen beiden Versuchen vermuten; eben jenen, nach denen sich die jeweiligen Stimuli-Paletten unterscheiden: Instrumente und Tonhöhen.

103 Grey (1976).
104 Lakatos (2000). S. 1434.
105 Melara et al. (1990).
106 Iverson (1995).

### 6.2.1 - Unterschiede hinsichtlich der Instrumente

Die offensichtlichsten Unterschiede liegen in den verwendeten Instrumenten. Als einzig gemeinsames Instrument kommt die Klarinette in beiden Versuchen zum Einsatz, was nur auf diese Stimuli reduziert, jedoch ebenfalls keine Gemeinsamkeiten im Vergleich der beiden MDS-Ergebnisse erkennen läßt. Dies wäre jedoch auch nicht zu erwarten, da einzelne Stimuli immer in Relation zu den anderen gesehen werden und diese wechselseitigen Beziehungen die MDS-Struktur bestimmen. Dementsprechend sind die Klarinetten-Stimuli in beiden Versuchen in jeweils ganz andere Kontexte eingebettet. Dagegen stammen alle übrigen Instrumentalklänge in beiden Versuchen aus unterschiedlichen Instrumenten, woraus grundlegende Unterschiede in beiden geometrischen Konstellationen X der MDS-Ergebnisse erwartet werden konnten. Mit der Beschränkung beider Versuche auf den Gebrauch von Blasinstrumenten konnte jedoch angenommen werden, daß die Anzahl und Rolle der Dimensionen ähnlich sein würde, was sich letztlich nur teilweise bestätigt.

Die schwächere Trennung zwischen Instrumenten bei Handel & Erickson (2004) läßt annehmen, daß ihre Kombination von Klarinetten-, Englischhorn- und Trompeten-Stimuli als ähnlicher in ihrer Klangfarbe bewertet wurden, als es für die Stimuli dieser Arbeit der Fall ist. Aus verschiedenen Versuchen zur Identifikation von Instrumentalklängen gibt es Erkenntnisse, daß Klänge der Klarinette besonders gut identifiziert werden können.[107] Zudem ist in den Ergebnissen von Strong & Clark und Berger zu erkennen, daß Horn- und Flötenklänge häufiger miteinander verwechselt werden. Diese Erkenntnisse sind in dieser Arbeit mit der auf zwei Dimensionen basierenden Trennung der Klarinette von den anderen Instrumenten sehr gut vereinbar. Aufgrund der im einzelnen unterschiedlichen Versuche läßt sich jedoch nur die allgemeine Annahme ableiten, daß die Klarinette klangliche Eigenschaften besitzt, welche ihre Klänge gut von anderen Instrumentalklängen unterscheiden lassen.

Bei Handel & Erickson (2004) hinterläßt diese gute Identifizierbarkeit der Klarinette jedoch einige Unklarheit, da sich diese Stimuli dort keineswegs deutlich von den anderen trennen. Es kommt über den verschiedenen Tonhöhen sogar zu Überkreuzungen der Klarinette mit dem Englischhorn und der Trompete. Handel & Erickson (2004) beabsichtigten mit ihrer Wahl der Klarinette und des Englischhorns eine gezielte Gegenüberstellung relativ ähnlicher Instrumente. Bei Strong & Clark treten zwischen diesen beiden Instrumenten zwar auch leichte Verwechslungen auf, doch läßt sich hieraus die punktuelle Überschneidung der Klarinette mit der Trompete kaum erklären. Handel & Erickson (2004) begründen diese Überschneidung bei $g^1$ mit dem Registerwechsel, dessen genaue Lage jedoch, wie in Abschnitt 2.4.1 gezeigt, entweder darunter (bei $d^1$) oder darüber (bei $d^2$) angesiedelt sein müßte. Alternativ kann man in dieser Lage auch den Einfluß der Überblasgrenze (bei $a^1$) erwarten.

Letztendlich lassen sich aus den Unterschieden der Instrumente die gravierenden Unterschiede in der Organisationsstruktur kaum erklären. Dies legt jedoch den Verdacht nahe, daß die beiden Haupteinflüsse für die jeweilige Struktur auf dieser Ebene einen aussagekräftigen, direkten Vergleich vielleicht gar nicht erst zulassen. Während sich demnach die deutliche Trennung von Instrumenten

107 vgl. Berger (1963), Saldanha & Corso (1964), Strong & Clark (1967), Brown (2000).

im Versuch dieser Arbeit zumindest ansatzweise mit den Ergebnissen aus Identifikationsversuchen vergleichen läßt, da eine Trennung nach Instrumenten und die Identifikation von Instrumenten durchaus vergleichbar sind, tragen derartige Vergleiche bei Handel & Erickson (2004) zu keiner weiteren Aufklärung bei. Es ist einleuchtend, daß der maßgebliche, strukturgebende Einfluß, sei es die Trennung nach Instrumenten oder dem Tonhöheneinfluß, eine zentrale Stellung im Gesamtergebnis einnimmt. Gleichzeitig treten weniger bedeutsame Faktoren in den Hintergrund und dadurch auch die mit ihnen verbundenen Vergleichsgrößen. Dies läßt sich exemplarisch an den zwei korrelierenden Vergleichsgrößen *spectral centroid* und *Tonhaltigkeit* bezogen auf beide Versuche verdeutlichen:

Die Vergleichsgröße *spectral centroid* ist (gerade im nicht-frequenznormierten Fall) eine Größe, die prinzipiell von der Tonhöhe abhängig sein kann, da mit unterschiedlichen Tonhöhen bzw. Grundtonfrequenzen sich ebenfalls das spektrale Zentrum verschiebt. Im Falle einer von der Tonhöhe geprägten MDS-Struktur, kann diese Größe, wie bei Handel & Erickson (2004) zu sehen, durchaus zur Erklärung der Beziehungen zwischen den Instrumenten und ihrer dimensionalen Abbildung herangezogen werden. Dieselbe Größe vermag es im Ergebnis dieser Arbeit (mit seiner durch Instrumententrennung geprägten Struktur) jedoch nicht die Trennung in Dimension 1 zu erklären. Dagegen läßt sich die Struktur sehr gut mit der Vergleichsgröße *Tonhaltigkeit* erklären, was vor allem dadurch begünstigt wird, daß diese Größe keine Abhängigkeit von der Tonhöhe besitzt. So werden, abhängig vom strukturgebenden Einfluß, unterschiedliche Rahmenbedingungen für weitere Einflüsse geschaffen. Entsprechend ihrer Ausprägung lassen sich diese Einflüsse daraufhin in manchen Fällen mit relevanten Vergleichsgrößen beschreiben.

### 6.2.2 - Unterschiede hinsichtlich der Tonhöhe

Wie im vorangegangenen Abschnitt zu sehen, können Unterschiede in den Instrumenten gerade wegen des Einflusses der Tonhöhe nicht zur ausreichenden Erklärung der unterschiedlichen Strukturprinzipien herangezogen werden. Die Versuche dieser Arbeit und von Handel & Erickson (2004) weisen zwar in beiden Fällen einen Einfluß der Tonhöhe auf, doch sind diese bekanntlich ganz unterschiedlich in ihrer Bedeutung für das MDS-Ergebnis. Die Frage der deutlichen Unterschiede bezüglich der Instrumententrennung läßt sich schließlich besser durch vorhandene Unterschiede bezüglich der Tonhöhen erklären. Es läßt sich zudem die Frage anfügen, woraus sich die grundsätzlich unterschiedlichen Rollen des Tonhöheneinflusses ergeben.

Klare und durchaus beabsichtigte Unterschiede gibt es bezüglich der konkreten Auswahl der Tonhöhen für den Versuch dieser Arbeit. Diese beziehen sich sowohl auf die jeweils gewählten absoluten Tonhöhen, als auch die aus ihnen resultierenden Intervalle. Es wurde in der Zielsetzung bereits begründet, daß die gewählten Tonhöhen eine differenziertere Wahl darstellen als bei Handel & Erickson (2004), wo eine Neigung zu oktavierten und eng verwandten Tonhöhen vorherrscht. Vermutlich sind diese Gründe auch für die entstandenen Unterschiede verantwortlich.

Anzeichen für die Relevanz dieser Vermutung liefert vor allem die Betrachtung der Ergebnisse von Miller & Carterette (1975). Sie untersuchten unter anderem den Einfluß der Tonhöhe in einem

MDS-Versuch zur Klangfarbe mit Tonkomplexen. Sie erhalten einen noch deutlich stärkeren Einfluß der Tonhöhe, als es bei Handel & Erickson (2004) der Fall ist. Ihre Auswahl synthetischer Klänge beschränkt sich ausschließlich auf die Verwendung von oktavierten Tonhöhen (200, 400 und 800 Hz, entsprechend etwa g, $g^1$ und $g^2$). Aufgrund des noch stärkeren Tonhöheneinflusses scheint die Ursache für diese Abhängigkeit bei Miller & Carterette möglicherweise mit der Oktavengleichheit der Töne in Verbindung zu stehen. In einer sehr ähnlichen und geringfügig schwächeren Ausprägung sind auch die Tonhöhen (c, g, $c^1$, $g^1$, $c^2$ und $g^2$) von Handel & Erickson (2004) hiervon betroffen. Beruhend auf der *Zweikomponententheorie* tragen zwei Eigenschaften zur Tonhöhenwahrnehmung bei: die *Tonigkeit* (Oktavengleichheit), eine pro Halbtonschritt charakteristische, qualitative ‚Färbung' eines Tons, welche sich zyklisch jede Oktave wiederholt und die *Helligkeit*, welche sich allein nach der physikalischen Grundtonfrequenz richtet und somit linear ansteigt.[108] Somit können Intervalle (größer einer Prime) sich entweder in beiden Komponenten unterscheiden, oder wie im Falle von Oktaven und ihren Vielfachen lediglich in der Helligkeit variieren. Demnach sind sich Oktaven wegen ihrer Gleichheit in der Tonigkeit in ihrer Tonempfindung deutlich ähnlicher, als alle anderen Intervalle. Bezogen auf den Versuch von Miller & Carterette bedeutet dies, daß sich die Tonhöhen nur in ihrer Helligkeit unterscheiden, womit der starke Tonhöheneinfluß ebenfalls nur in Unterschieden der linearen Komponente Helligkeit fundiert sein kann. Die Helligkeit ist laut de la Motte-Haber für die Hörwahrnehmung wichtiger als die Tonigkeit[109], woraus durch die Abwesenheit von Tonigkeitsunterschieden ihr Einfluß intensiviert werden dürfte.

Für Handel & Erickson (2004) kann man aufgrund der Oktavierung der Tonhöhen c und g auch auf einen starken Einfluß der Helligkeit schließen. Hinzu kommt, daß die in der Tonigkeit unterschiedlichen Tonhöhen c und g trotzdem eng miteinander verwandt sind. Intervallverhältnisse einer Quinte werden, wie auch Oktaven, als ‚reine Intervalle' angesehen, da sie von allen Intervallen die engste Tonverwandschaft besitzen. Physikalisch läßt sich dies damit erklären, daß in einer harmonischen Obertonstruktur der erste in der Tonigkeit abweichende Teilton die Quinte (Duodezime) darstellt. Daher besteht ein harmonischer Klang überwiegend aus Vielfachen der Oktave und Quinte des Grundtons. In geringerem Umfang als bei den Oktaven sind bei Handel & Erickson (2004) somit selbst die Tonhöhen c und g sehr ähnlich, was in Intervallen bestehend aus diesen zwei Tonigkeiten einen stärkeren Einfluß der Komponente Helligkeit wie bei Miller & Carterette erwarten läßt.

Die in dieser Arbeit verwendeten Tonhöhen (g, $c^1$ und $h^1$) weisen dagegen keine Oktavbeziehungen zueinander auf, womit sie sich in allen Fällen in beiden Komponenten Tonigkeit und Helligkeit unterscheiden. Zwar wäre auch hier zu argumentieren, daß die Tonhöhen g und $c^1$ vom Teiltongehalt sehr viele Ähnlichkeiten aufweisen. Nur handelt es sich in diesem konkreten Fall effektiv um ein Quartverhältnis, welches zudem nicht weiter oktaviert wird. Außerdem ergeben alle möglichen Kombinationen der Tonhöhen sich ausschließende Intervalle (reine Quarte, große Septime und große Dezime). Die letzten beiden Intervalle weisen einen relativ geringen Konsonanzgrad auf, womit sie sich von ‚reinen Intervallen' (Oktave, Quinte) klar unterscheiden. Die Quarte (g–$c^1$) gilt hingegen

108 vgl. v. Hornbostel (1926), Bachem (1950), Shepard (1982), sowie Singh & Hirsh (1992), Marozeau et al. (2003).
109 de la Motte-Haber (1996). S. 464.

ebenfalls als ‚reines Intervall', woraus sich die Inkonsistenzen für dieses Intervall im MDS-Ergebnis ferner erklären könnten.

Diese Erkenntnisse liefern Anzeichen dafür, daß der starke Tonhöheneinfluß bei Miller & Carterette, sowie Handel & Erickson (2004) durch eine zu starke Fokussierung auf die Komponente Helligkeit begünstigt wird. Diese Annahme wird zudem durch die Ergebnisse von Marozeau et al. (2003) bestärkt. Sie untersuchten ausschließlich den Einfluß der Tonhöhe auf MDS-Versuche zur Klangfarbe. In ihrer Dimension 4 läßt sich ein Tonhöheneinfluß nachweisen, der jedoch keine strukturgebende Stellung einnimmt. Sie fassen zusammen, daß Ähnlichkeit der Klangfarbe in einem begrenzten Maß von Unterschieden der Grundtonfrequenz bzw. der Tonhöhe abhängt und daß bei ihnen jedoch in allen Fällen tonhöhen-unabhängige Faktoren stärker ins Gewicht fallen als tonhöhen-abhängige.[110] Bezüglich der von ihnen gewählten Tonhöhen (h, cis$^1$ und b$^1$) unterscheiden auch diese sich in den beiden Komponenten Tonigkeit und Helligkeit. Dies deutet möglicherweise darauf hin, daß bei der Bewertung von Klangfarben vorhandene Unterschiede in der Tonigkeit es erleichtern könnten, von Tonhöhenunterschieden abzusehen bzw. einem stärkeren Einfluß der Komponente Helligkeit zu begegnen.

Diesbezüglich entsteht jedoch ein Widerspruch zu den Ergebnissen von Erickson (2003), der an dieser Stelle daher noch diskutiert werden sollte. Erickson führte einen MDS-Versuch mit unterschiedlichen Tonhöhen (a, c$^1$, g$^1$, h$^1$, f$^2$ und a$^2$) durch, in diesem Fall jedoch mit Klängen von Singstimmen als Stimuli. Trotz überwiegender Unterschiede in beiden Tonhöhenkomponenten Tonigkeit und Helligkeit kommt es in ihrem Ergebnis jedoch zu einem vergleichbaren starken Einfluß der Tonhöhe wie bei Handel & Erickson (2004). Dieses würde also der Annahme einer dominanten Bedeutung der Komponente Helligkeit für den Tonhöheneinfluß ein Stück weit widersprechen, weil in diesem Fall auch mehrere unterschiedliche Tonigkeiten vorliegen. Es läßt sich jedoch annehmen, daß der große Tonhöheneinfluß in diesem konkreten Fall von einem anderen Kontext abhängig ist. Als Klangquellen für die Stimuli dienten jeweils zwei unterschiedliche Sopran- und Mezzosopranstimmen. Es ist anzunehmen, daß Vergleiche der Ähnlichkeit zwischen Singstimmen derselben oder vergleichbaren Gesangskategorie gegenüber Vergleichen zwischen Instrumenten allgemein zu höheren Bewertungen in der Ähnlichkeit führen. Anhand dieser Annahme kann argumentiert werden, daß aufgrund der hohen Ähnlichkeit der Stimuli hinsichtlich ihrer Klangfarbe bestehende Unterschiede in der Tonhöhe bzw. der Komponente Helligkeit mehr in den Mittelpunkt des Wahrnehmungsfelds rücken und somit auch einen größeren Einfluß auf die Bewertungen ausüben. Anders ließe sich dieser Widerspruch, bezüglich der Begünstigung eines Tonhöheneinflusses durch Vorliegen gleicher Tonigkeiten, nicht aufklären. Miller & Carterette äußern sich diesbezüglich: „[S]ubjects may try to minimize their efforts in the experiment by responding to only the very obvious distinctions present."[111] Gleichzeitig liefert diese Erkenntnis jedoch auch Anzeichen dafür, daß der Einfluß der Tonhöhe auf die MDS-Struktur sehr wohl von der Zusammensetzung bzw. auch der Quellen der Stimuli abhängig ist, wie bereits im Vergleich der Unterschiede hinsichtlich der Instrumente angemerkt wurde.

110 Marozeau et al. (2003). S. 2950.
111 Miller & Carterette (1975). S. 718.

Als weitere Unterschiede, die zur unterschiedlichen Bedeutung der Tonhöhe für beide MDS-Ergebnisse beitragen, können noch die Anzahl der verwendeten Tonhöhen, sowie ihre Ausdehnung im Tonraum bzw. ihre Intervallgrößen genannt werden. So könnte die Anzahl der Tonhöhen bei Handel & Erickson (2004) bzw. Erickson (2003) mit fünf, respektive sechs Tonhöhen einen Einfluß der Tonhöhe begünstigen, während in dieser Arbeit und bei Marozeau et al. (2003) mit jeweils drei Tonhöhen, sich der Tonhöheneinfluß in Grenzen hält. Gleichzeitig ließe sich dieser Logik nach bei Miller & Carterette mit ebenfalls drei Tonhöhen der noch viel stärkere Einfluß der Tonhöhe jedoch nicht begründen. Dies ist ein Indiz dafür, daß der Einfluß der Komponente Helligkeit hier doch eine wichtigere Rolle spielen könnte. Diese Logik läßt sich auch auf den Umfang der Tonhöhen anwenden. Die Tonhöhen von Miller & Carterette (1975), Erickson (2003) und Handel & Erickson (2004) umfassen zwei bis zweiundhalb Oktaven. Dagegen betragen die Tonhöhenumfänge in dieser Arbeit und bei Marozeau et al. (2003) in etwa eine Oktave. Handel & Erickson (2001) stellen bezüglich eines Versuchs mit der Zuordnung von Klängen eines gleichen bzw. anderen Instruments fest, daß es bei Intervallen einer Oktave und größer zu vielen Fehlbewertungen von Klangpaaren gleicher Instrumente kommt. Sie begründen dies, mit der durch größere Tonhöhenunterschiede entstehenden klanglichen Veränderung, die eine nachlassende direkte Identifizierbarkeit von Klängen gleicher Instrumente verursacht. Mit den zahlreichen Tonhöhen bei Handel & Erickson (2004) und Erickson (2003) resultieren aus allen möglichen Tonhöhenkombinationen auch Intervalle kleiner einer Oktave. Die Anzahl größerer Intervalle könnte aber ausgereicht haben, um statt von der Tonhöhe unabhängigen Unterschieden in der Klangfarbe, welche nicht mehr leicht zu extrahieren waren, primär nach tonhöhenabhängigen Unterschieden bewertet zu haben. Die MDS-Struktur orientiert sich infolgedessen nach der Tonhöhenentwicklung. Marozeau et al. (2003) stellen zudem fest, daß sich für ihr kleinstes Intervall h–cis$^1$ kein Tonhöheneinfluß nachweisen läßt, womit sich der Tonhöheneinfluß nur im deutlich größeren Intervall h–b$^1$ abbildet.[112] Ähnliches könnte auch über die Abbildung des Tonhöheneinflusses in dieser Arbeit gesagt werden, indem der Tonhöheneinfluß sich nur für die größeren Intervalle g–h$^1$ und c$^1$–h$^1$ klar abbilden läßt, während es beim kleinen Intervall g–c$^1$ bekanntlich zu inkonsistenten Tendenzen kommt. Es sei betont, daß die großen Intervalle in all diesen Fällen annähernd eine Oktave betragen. Die Rolle des verwendeten Tonumfangs scheint daher kein zu vernachlässigender Faktor zu sein.

Der letzte die Tonhöhe betreffende zu berücksichtigende Faktor bezieht sich auf die musikalische Vorbildung der VP. In Abschnitt 1.2.1 wurde bereits angemerkt, daß die meisten MDS-Studien mit einer Tonhöhe keinen Unterschied zwischen den Bewertungen musikalisch Vorgebildeter und Laien finden. Es wurde jedoch auch vorweggenommen, daß die musikalische Vorbildung besonders im Hinblick auf MDS-Studien mit unterschiedlichen Tonhöhen berücksichtigt werden müßte. Gerade in Anbetracht der Komplexität der an die VP gestellten Bewertungsaufgabe, Klangfarbenunterschiede unabhängig von Tonhöhenunterschieden zu bewerten, könnten Unterschiede in der musikalischen Vorbildung zu grundsätzlich unterschiedlichen MDS-Ergebnissen führen. Bei Handel & Erickson (2004) nahmen musikalische Laien am Versuch teil, während am Versuch dieser Arbeit ausschließlich musikalisch vorgebildete VP beteiligt waren. Daher muß hinterfragt werden, ob die

112 Marozeau et al. (2003). S. 2955.

VP dieser Arbeit durch ihre musikalische Vorbildung befähigter waren, zwischen Unterschieden der Klangfarbe und denen der Tonhöhe zu abstrahieren. Könnte dadurch ein geringerer Einfluß der Tonhöhe in der MDS-Struktur begründet werden?

Pitt (1994) führte einen Versuch mit Paarvergleichen durch, orientiert an einem Kreuzdesign mit vier Konditionen, bei denen zwei Konditionen mit isolierten Tonhöhen- bzw. Klangfarbenänderungen und zwei Konditionen mit korrelierten Änderungen in beiden Bereichen verbunden waren. Die Versuchsergebnisse zeigen, daß musikalische Vorbildung die perzeptive Trennung von Klangfarben- und Tonhöhenänderungen signifikant verbessert. Es sei angemerkt, daß Pitts Tonhöhenunterschied einer übermäßigen Quarte (Tritonus) entspricht und somit in beiden Komponenten Tonigkeit und Helligkeit Unterschiede aufweist, sowie einen sehr geringen Konsonanzgrad besitzt. Krumhansl & Iverson (1992) führten einen ähnlichen Versuch mit ausschließlich musikalisch vorgebildeten VP durch. Sie entdeckten eine Interaktion zwischen Klangfarben- und Tonhöhenverarbeitung in signifikanten Änderungen der Reaktions- bzw. Bewertungszeit in den verschiedenen Konditionen. Abgesehen von diesen zeitlichen Effekten hatten alle Konditionen keinen Einfluß auf die Rate der korrekten Antworten. Dies bedeutet, daß für ein Versuchsmodell ohne Notwendigkeit von schnellen Beantwortungen generell kein negativer Effekt auf die Bewertungsqualität erwartet werden dürfte. Somit ist anzunehmen, daß für einen Versuch dieser Art musikalisch Vorgebildete in der Regel als VP bevorzugt werden sollten. Die Wahl von Handel & Erickson (2004) für musikalische Laien als VP erfolgte, nachdem in einem vergleichbaren Versuch bei Erickson (2003) kein Unterschied für die musikalische Vorbildung gefunden wurde. Für letzteren, bereits erwähnten Versuch mit Singstimmen wurde bereits angemerkt, daß die Stimuli eine hohe Ähnlichkeit aufweisen, so daß es sein könnte, daß selbst musikalisch Vorgebildete aufgrund der geringeren Unterschiedlichkeit der Stimuli ihre Bewertungen nach tonhöhenabhängigen Faktoren richteten. Auch Miller & Carterette (1975) untersuchten in ihrem MDS-Versuch mit unterschiedlichen Tonhöhen den Einfluß der musikalischen Vorbildung und fanden ebenso keinen Unterschied zwischen den beiden Gruppen. Bei ihnen sind auch Bewertungen von musikalisch vorgebildeten VP stark von Tonhöhenunterschieden beeinflußt, was jedoch im alleinigen Vorhandensein von Unterschieden in der Helligkeit begründet liegen kann. Bezüglich Pitts Feststellungen muß zudem noch erwähnt werden, daß sie sich nicht ohne weiteres auf MDS-Versuche beziehen lassen. In Pitts Versuch waren sich die VP nicht über den Gegenstand der Untersuchung bewußt, d.h. sie wußten nicht welche interagierenden Faktoren vorhanden waren. Im MDS-Versuch dieser Arbeit wurden die VP dagegen ausdrücklich aufgefordert, Unterschiede in der Tonhöhe zu ignorieren. Diese bewußte Aufforderung müßte jedoch, zumindest in geringem Maße, ein effektives Ignorieren von Tonhöhenunterschieden verbessern können.

### 6.2.3 - Weitere Unterschiede

Bezüglich der musikalischen Vorbildung kann zusätzlich eine weiterer Effekt angenommen werden. Dieser bezieht sich auf unterschiedliche Bewertungsansätze von VP beim Vergleich der Ähnlichkeit. Musikalische Vorbildung kann hier wesentlich komplexere und unter Umständen auch von zu vielen rationalen Überlegungen geprägte Bewertungskriterien zu Tage fördern, im Vergleich zu einer relativ

naiven und eventuell sogar sensorisch geprägteren Bewertung durch einen musikalischen Laien. Kendall et al. (1998) nennen diesbezüglich die Möglichkeit, daß eine VP in ihrer Bewertung als erstes die Instrumente identifiziert und infolge anhand dieses Bias den Vergleich eines Stimuli-Paars vornimmt.[113] Letztendlich würden so eher konzeptionelle, verallgemeinerte Unterschiede zwischen Stimuli eines Instruments bewertet, als die tatsächlichen, perzeptiven Unterschiede zwischen voneinander unabhängigen Stimuli. Für das Resultat dieser Arbeit wäre diese Möglichkeit auch anzunehmen, da es zu einer sehr klaren Trennung aller Instrumente kommt. Die musikalische Vorbildung würde mit einer eventuell besseren Fähigkeit der Instrumentenidentifikation dahingehend die oben genannte Möglichkeit begünstigen.

Auch die semantische Bedeutung des Begriffs *Klangfarbe* ist ein bedeutender Faktor für unterschiedliche Bewertungsansätze. Es gibt bekanntlich keine Definition, welche eine hinreichende Beschreibung und Ein- bzw. Abgrenzung der Klangfarbe gewährleistet. Außerdem existieren unterschiedliche und widersprüchliche Verständnisse über die Klangfarbe. Man kann in der Versuchsanweisung lediglich versuchen, eine Spezifizierung des Begriffs vorzunehmen; bei Handel & Erickson (2004) ohne Erwähnung des Begriffs *timbre* geschehen, während in dieser Arbeit *Klangfarbe* ausdrücklich genannt wurde. Diesbezüglich sind zusätzlich noch auf sprachlicher Inkompatibilität beruhende semantische Unterschiede denkbar. Kann man demnach die Begriffe *timbre* und *Klangfarbe* ohne weiteres gleichsetzen?

Als Letztes ist nicht auszuschließen, daß die gegenüber dem Versuch von Handel & Erickson (2004) vorgenommenen Änderungen bezüglich Stimuli-Dauer und -lautstärke, sowie die stereophone Wiedergabe, auch einen Anteil an der unterschiedlichen dimensionalen Struktur haben. Ihren Einfluß kann man jedoch mit den verfügbaren Mitteln nicht abschätzen, da es aufgrund der bestehenden Unterschiede zwischen den Versuchen keine Vergleichsgrundlage gibt. Die einzige Möglichkeit wäre gewesen, einen dieser Faktoren mittels einer Kontrollgruppe als eigene unabhängige Variable in den Versuchsplan zu implementieren. Dies wäre im Umfang dieser Magisterarbeit jedoch nicht möglich gewesen. Die im Allgemeinen vergleichbaren Ergebnisse aus den bereits genannten MDS-Studien zur Klangfarbe, in denen bezüglich der Standardisierung teilweise erhebliche Unterschiede existieren, sprechen jedoch für die relativ sichere Annahme einer marginalen Bedeutung für das MDS-Ergebnis.

113 Kendall et al. (1998). S. 354.

## 6.3 - Zusammenfassung und Ausblick

Wie in vielen Studien zuvor bestätigt sich in dieser Arbeit abermals, daß sowohl spektrale als auch temporale Eigenschaften als Einflüsse für die dimensionale Struktur des MDS-Ergebnisses in Frage kommen. Spektrale Eigenschaften in Gestalt der psychoakustischen Größe *Tonhaltigkeit* lassen sich gut mit der Struktur in Dimension 1 vereinbaren. Dies deutet auf eine vorliegende Unterscheidung zwischen Stimuli mit größeren und geringeren Geräuschanteilen. Dagegen könnten temporale Eigenschaften, in Form von Unterschieden in der Einschwingdauer der Klänge, die Verteilung der Stimuli in Dimension 2 bestimmt haben. Gleichzeitig müssen jedoch auch wahrnehmbare, subjektive Unterschiede im Klang als Einfluß hierfür angenommen werden, welche eventuell mit der erwähnten ‚Nähe' bzw. ‚Direktheit' des Klarinettenklangs zusammenhängen. In Dimension 3 tritt ein Einfluß der Tonhöhe hinzu, welcher sich auch in allen genannten MDS-Studien mit Stimuli unterschiedlicher Tonhöhen nachweisen läßt. Es liegen jedoch teils deutliche Abweichungen von einer idealen Abbildung der Tonhöhe vor. Hinzu kommt, daß Ergebnisse in Dimension 3 als nicht empirisch abgesichert gelten müssen. Bezüglich der Verwendung der gewählten Vergleichsgrößen zur quantitativen Überprüfung der MDS-Struktur zeigt sich, daß einige von Handel & Erickson (2004) verwendete Größen sowie im beschränkten Maße auch psychoakustische Größen zur Beschreibung von Instrumentalklängen unterschiedlicher Tonhöhe herangezogen werden können.

Die in dieser Arbeit vorliegende MDS-Struktur orientiert sich an einer deutlichen räumlichen Trennung einzelner Instrumente. Klangvariationen eines selben Instruments (Horn), werden hierbei nicht berücksichtigt. Dies deutet auf eine Kategorisierung der Stimuli entsprechend der Instrumente. Jedoch treten auch keine lokal kompakten Cluster dieser Stimuli-Gruppen auf, wie von Handel & Erickson (2004) für den Fall einer alleinigen Identifikation der Instrumente als mögliches Strukturprinzip postuliert wird (siehe Abschnitt 1.2.2). Es findet vielmehr eine ‚Auffächerung' dieser auf Dimensionen 1 und 2 basierten Trennung entlang der Tonhöhe in Dimension 3 statt. Dies weist auf eine stärkere Gewichtung inter-instrumentaler Unterschiede hin, wogegen auf intra-instrumentaler Ebene die Tonhöhe verstärkt zu Unterschieden in der Klangfarbe beiträgt.

Handel & Erickson (2004) suchten dem Titel ihres Artikels[114] nach eine Identifikation von Instrumenten bzw. Klangquellen. Aus ihrem maßgeblich vom Tonhöheneinfluß geprägten Ergebnis, welches verfehlt eine Identifikation von Instrumenten widerzuspiegeln, folgern sie, daß ein MDS-Versuch mit Vergleichen zwischen nur zwei Klängen zu keiner Identifikation von Instrumenten bzw. Klangquellen führen kann. Nach ihnen stützt sich, wie eingangs erwähnt, die Identifikation von Instrumenten auf instrumentenspezifische Kenntnisse bezüglich eines ‚Satzes an Klangfarben', welcher durch eine solch geringe Zahl von Klängen nicht vermittelt werden kann. Im Ergebnis dieser Arbeit wird diese Annahme jedoch klar widerlegt, da ohne weiteres eine Trennung nach Instrumenten erreicht wird. Als Folge kann aus keiner der beiden Versuche eine universelle Anwendbarkeit auf MDS-Versuche mit unterschiedlichen Tonhöhen abgeleitet werden. Vielmehr deutet alles auf eine starke Kontextabhängigkeit, aus der sich die Unterschiede ergeben. Die grundlegenden Unterschiede in den

114 „Sound Source Identification: The Possible Role of Timbre Transformations".

beiden Versuchen beruhen dem Anschein nach auf Kontextunterschieden aus den gewählten Tonhöhen, der Instrumente, sowie auch auf der Auswahl von VP und ihrer musikalischen Vorbildung.

Um die Relevanz der Unterschiede bezüglich der Tonhöhe im einzelnen überprüfen zu können, würde sich eine umfassende Reproduktion des Versuchs mit einem multifaktoriellen Plan anbieten. So könnten neben einer zu dieser Arbeit nahezu identischen Versuchsdurchführung als Kontrollbedingung, gleichzeitig zwei weitere Bedingungen untersucht werden: einerseits mit mehr Tonhöhen bzw. einem größeren Tonhöhenumfang, andererseits mit dem Verzicht bzw. der sparsamen Nutzung von Tonhöhen mit unterschiedlicher Tonigkeit. Zur Erfassung unterschiedlicher Bewertungshaltungen wie zum Beispiel aus Unterschieden in der musikalischen Vorbildung, empfehlen sich MDS-Algorithmen mit Berücksichtigung individueller Unterschiede. Mit einigem Erfolg wurde hierzu bezüglich der Klangfarbenforschung bereits das CLASCAL-Verfahren eingesetzt. Dieses kann neben dem Aufdecken von unterschiedlichen Bewertungshaltungen, zudem auch nicht in den Dimensionen abbildbare Eigenarten von Instrumenten mittels sogenannter Spezifitäten (*specificities*) ausdrücken.[115]

In dieser Arbeit kann des weiteren erkannt werden, daß die Bedeutung dieser instrumentenspezifischen Eigenarten keineswegs unterschätzt werden darf. Denn die stärksten korrelierenden Vergleichsgrößen vermögen es bei weitem nicht die MDS-Struktur hundertprozentig zu erklären. Es sei ins Gedächtnis gerufen, daß spezifische Eigenschaften einzelner Instrumente im Gesamtkontext additiv bzw. interagierend zur Unterschiedlichkeit beigetragen haben könnten. Dies setzt eine komplexere Auswertung bzw. Bewertung der Stimuli-Paare voraus, als mit einer Skalierung in wenigen Dimensionen ausgedrückt werden könnte. Weiterhin ist eine kategoriale Beurteilung in manchen Fällen denkbar, wie im Falle von ‚langer' und ‚kurzer' Einschwingdauer.

Außerdem deutet eine kategoriale Beurteilung darauf hin, daß es nicht zwangsläufig zu rein perzeptiven Bewertungen von physikalisch-akustischen Unterschieden auf einer niedrigen Ebene der Informationsverarbeitung kommt, sondern es zu einem Beitrag von bzw. zu Interaktion mit kognitiven Prozessen kommt. Besonders die vorliegende Trennung bzw. Identifikation der Instrumente muß hierfür als Beleg gesehen werden. So kann letztlich mit MDS-Versuchen dieser Art kein streng psychophysisches Experiment angenommen werden, da kognitive Prozesse in der Beurteilung bei weitem nicht ausgeschlossen werden können. In der Regel kommt es nach de la Motte-Haber zu einer Interaktion zwischen Verarbeitungsprozessen, welche „durch Kenntnisse gesteuert (top-down) und durch Reize induziert (bottom-up)" sind.[116]

Mit Blick auf den möglichen Einfluß kategorialer Urteile erscheint ein Beharren auf die Skalierbarkeit von wirksamen Einflüssen entsprechend des Intervallskalenniveaus nicht sinnvoll. Mit einer kategorialen Trennung sind oftmals nur Skalenniveaus einer Nominal- oder Ordinalskala zu erreichen. Diese bringen einen statistisch relevanten Nachteil mit sich, indem sie sich weniger zur quantitativen Überprüfung mittels Korrelationsanalysen mit den intervallskalierten MDS-Dimensionen eignen. Mit

115 vgl. McAdams et al. (1995) und Marozeau et al. (2003).
116 de la Motte-Haber (1996). S. 462.

einer Abkehr von der strengen Skalierbarkeit der Klangfarbe wendet man sich jedoch letztlich einem Ansatz hin, welcher der tatsächlichen menschlichen Hörwahrnehmung – im alltäglichen Umgang mit musikalischen Klängen und nicht innerhalb einer künstlich kontrollierten Laborsituation – eher entspricht.

Es kann vieles über Klangfarbe gesagt werden: Sie ist multidimensional. Sie läßt sich mit einer endlos langen Liste von physikalisch-akustischen Eigenschaften beschreiben. Sie kann gleichzeitig zur Unterscheidung von Klängen unterschiedlicher Instrumente, zweier Instrumente der gleichen Gattung, oder sogar ein und desselben Instruments dienen. Nur eines bleibt auch fortwährend ungeklärt, nämlich welche und wie viele dieser Eigenschaften in einem konkreten Fall letztlich zur Geltung kommen. Bezogen auf diese Unklarheit in MDS-Untersuchungen äußert sich Sixtl, daß „man bei höher dimensioniertem Reizmaterial damit rechnen [muß], daß die [VP] nur eine begrenzte Anzahl von Merkmalen beachten kann [...]“[117]. Dies weist auf die in dieser Arbeit klar erkenntlich gewordene wesentliche Eigenschaft der Klangfarbe: ihre kontextabhängige Natur. Weiterhin ist auch fraglich, ob sich Klangfarbe überhaupt absolut ausdrücken läßt, wenn die Kontextabhängigkeit doch gerade auf relativen Bezügen beruht. Zudem beschränkt sich bekanntlich das Feld des Kontexts bei weitem nicht nur auf rein perzeptive Faktoren aus der sensorischen Verarbeitung von physikalisch-akustischen Eigenschaften.

Für die Zukunft empfehlen sich diesbezüglich besonders auch Untersuchungen hinsichtlich der Informationsverarbeitung der Klangfarbe bzw. auch den dafür in Frage kommenden Wahrnehmungsmodellen. Dies ist nicht zwingend mit einem Verzicht auf MDS als Analyseverfahren verbunden. Jedoch sollte bei Verwendung von MDS das Augenmerk auf eine qualitative, inhaltliche Analyse ihrer Struktur gerichtet werden, wie auch in dieser Arbeit mit einigem Erfolg gezeigt werden konnte. Als Wahrnehmungsmodelle sollten diejenigen in Frage kommen, welche weder an einem überwiegend operationistischen Reiz-Empfindungs-Ansatz (*Psychophysik*), noch einem kognitive Einflüsse ausschließenden Ansatz (*Gestalttheorie*) orientiert sind. Ein geeignetes Modell würde demnach eine Interaktion zwischen perzeptiven und kognitiven Prozessen erlauben, wenn nicht sogar voraussetzen, nachdem Reize innerhalb eines Gesamtkontexts wahrgenommen werden. Als Beispiele würden die *Bezugssystemtheorie* oder auch die *ökologische Wahrnehmungstheorie* in Frage kommen.[118]

Die vorangegangene Diskussion verdeutlicht die Komplexität des Sachverhalts *Klangfarbe*. Ein Ziel von Wissenschaft ist immer Definitionen zu finden, welche möglichst in Kürze einen Sachverhalt präzise und umfassend eingrenzen können. Für den facettenreichen Sachverhalt Klangfarbe zeigt sich, daß eine Festlegung auf eine derartige Definition nicht möglich ist, ohne daß damit ein Ausschluß von (in bestimmten Situationen) relevanten Eigenschaften einhergeht. Mit einem mehr auf inhaltliche, qualitative Analyse gerichteten Ansatz entfernt man sich weiter von dieser klaren Definierbarkeit, indem weniger mit quantitativen Werkzeugen gearbeitet wird, sondern man sich einer deskriptiven,

117 Sixtl (1982). S. 352.
118 vgl. Hellbrück (1993). S. 178-180 und 203-205.

hypothetischen Ebene zuwendet. Dies ist allerdings gegenüber der Psychophysik mit dem Nachteil nicht vorweisbarer quantifizierbarer Ergebnisse behaftet. Jedoch wird selbst mit der Formulierung von psychometrischen Funktionen zwischen den Ein- und Ausgängen eines Wahrnehmungssystems auch nur eine *black box*[119] der Unklarheit umgangen, in der die hier interessanten, latent wirksamen Prozesse verborgen bleiben. Die wenig faßbare Natur der Klangfarbe wird umso komplexer, wenn statt isolierter Klänge auch ihre Rolle in musikalischen Phrasen betrachtet wird. Es ist unter Umständen daher sinnvoller zu untersuchen, *wie* die Klangfarbe wahrgenommen wird, bevor man sich am Definieren des *was*-die-Klangfarbe-ist zu lange aufhält. Am Ende steht fest: Mit ihrer immanenten Kontextabhängigkeit wird die Klangfarbe wohl immer die mehrdeutigste und am schlechtesten zu definierende Komponente für die Wahrnehmung von Klängen bleiben.

## Danksagung

Für ihr Engagement und ihre Mitwirkung bedanke ich mich herzlich bei folgenden Personen. Ohne sie wäre die Aufzeichnung und Produktion der Stimuli sowie die Durchführung des Versuchs nicht möglich gewesen.

Beate-Gabriela Schmitt
Fabian Gawlik
Folkmar Hein
Johannes Lamotke
Kaori Sako
Robin Seiler

119 Hellbrück (1993). S. 173.

# LITERATURVERZEICHNIS


ANSI. (1973). *American national standard: Psychoacoustical terminology*. S3.20-1973. New York: American National Standards Institute.

Bachem, A. (1950). „Tone Height and Tone Chroma as Two Different Pitch Qualities“. *Acta Psychologica*, 7. S. 80-99.

Berger, Kenneth. W. (1964). „Some factors in the recognition of timbre“. *Journal of the Acoustical Society of America*, 36. S. 1888-1891.

Borg, Ingwer und Groenen, Patrick J. F. (2005). *Modern Multidimensional Scaling: Theory and Applications*. 2. Auflage. New York: Springer-Verlag.

Bregman, Albert S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, Massachusetts: MIT Press.

Brown, Judith C., Houix, O. und McAdams, S. (2001). „Feature dependence in the automatic identification of musical woodwind instruments“. *Journal of the Acoustical Society of America*, 109. S. 1064-1072.

Butler, David. (1992). *The Musician's Guide to Perception and Cognition of Music*. New York: Schirmer Books (Macmillan).

de la Motte-Haber, Helga. (1996). *Handbuch der Musikpsychologie*. 2. Auflage. Laaber: Laaber Verlag.

Dickreiter, Michael. (1997). *Handbuch der Tonstudiotechnik*. 6. Auflage. München: Saur.

Dowling, W. Jay und Harwood, Dane L. (1986). *Music Cognition*. London: Academic Press.

Erickson, Molly L. (2003). „Dissimilarity and the Classification of Female Singing Voices: A Preliminary Study“. *Journal of Voice*, 17, No. 2. S. 195-206.

Grey, John M. (1977). „Multidimensional perceptual scaling of musical timbres“. *Journal of the Acoustical Society of America*, 61. S. 1270-1277.

Grey, John M. und Moorer, James A. (1977). „Perceptual evaluations of synthesized musical instrument tones“. *Journal of the Acoustical Society of America*, 62. S. 454-462.

Hajda, John M., Kendall, R. A., Carterette, E. C. und Harshberger, Michael L. (1997). „Methodological issues in timbre research“. *Perception and Cognition of Music*. Hrsgs.: Deliège, Irène und Sloboda, John. Hove, East Sussex: Psychology Press. S. 253-306.

Handel, Stephen und Erickson, Molly L. (2001). „A rule of thumb: The bandwidth for timbre invariance is one octave“. *Music Perception*, 19. S. 121-126.

Handel, Stephen und Erickson, Molly L. (2004). „Sound Source Identification: The Possible Role of Timbre Transformations“. *Music Perception*, 21. S. 587-610.

Hellbrück, Jürgen. (1993). *Hören: Physiologie, Psychologie und Pathologie*. Göttingen: Hogrefe Verlag.

Iverson, Paul. (1995). „Auditory stream segregation by musical timbre: Effects of static and dynamic acoustic attributes“. *Journal of Experimental Psychology: Human Perception and Performance*, 21. S. 751-763.

Iverson, Paul und Krumhansl, Carol L. (1993). „Isolating the dynamic attributes of musical timbre“. *Journal of the Acoustical Society of America*, 94, 2595-2603.

Kendall, Roger A., Carterette, E. C. und Hajda, J. M. (1998). „Perceptual and acoustical features of natural and synthetic orchestral instrument tones“. *Music Perception*, 16. S. 327-364.

Krumhansl, Carol L. und Iverson, Paul. (1992). „Perceptual Interactions Between Musical Pitch and Timbre“. *Journal of Experimental Psychology: Human Perception and Performance*, 18. S. 739-751.

Kruskal, Joseph B. und Wish, Myron. (1978). *Multidimensional Scaling*. Beverly Hills/London: Sage Publications.

Lakatos, Stephen. (2000). „A common perceptual space for harmonic and percussive timbres“. *Perception and Psychophysics*, 62. S. 1429-1439.

Licklider, Joseph C. R. (1951). „Basic Correlates of the Auditory Stimulus“. *Handbook of Experimental Psychology*. Hrsg.: Stevens, S. S. New York: Wiley & Sons. S. 985-1039.

Marozeau, Jeremy., de Cheveigné, A., McAdams, S. und Winsberg, S. (2003). „The dependency of timbre on fundamental frequency“. *Journal of the Acoustical Society of America*, 114. S. 2946-2957.

McAdams, Stephen, Winsberg, S., Donnadieu, S., De Soete, G. und Krimphoff, J. (1995). „Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes“. *Psychological Research*, 58. S. 177-192.

Melara, Robert D. und Marks, Lawrence E. (1990). „Interaction among auditory dimensions: Timbre, pitch, and loudness“. *Perception and Psychophysics*, 48. S. 169-178.

Meyer, Jürgen. (1999). *Akustik und musikalische Aufführungspraxis*. 4. Auflage. Frankfurt am Main: Verlag Erwin Bochinsky.

Miller, James R. und Carterette, Edward C. (1975). „Perceptual space for musical structures“. *Journal of the Acoustical Society of America*, 58. S. 711-720.

Orth, Bernhard. (1979). *Rating-Verfahren und Grössenschätz-Methode: Skalenniveau und funktionale Zusammenhänge zwischen Skalen*. Dissertation. Kiel: Mathematisch-Naturwissenschaftliche Fakultät, Christian-Albrechts-Universität.

Pitt, Mark A. (1994). „Perception of pitch and timbre by musically trained and untrained listeners“. *Journal of Experimental Psychology: Human Perception and Performance*, 20. S. 976-986.

Plomp, Reinier. (1976). *Aspects of Tone Sensation: A Psychophysical Study*. London: Academic Press.

Reuter, Christoph. (2005). „Klangfarbe: Beziehungen zur Tonhöhe und Lautstärke – Merkmale der Mehrdimensionalität“. *Musikpsychologie. Handbuch der Systematischen Musikwissenschaft*, Bd. 3. Hrsgs.: de la Motte-Haber, Helga und Rötter, Günther. Laaber: Laaber Verlag. S. 250-262.

Rimsky-Korsakov, Nikolay. (1964). *Principles of Orchestration*. New York: Dover Publications.

Saldanha, E. L. und Corso, John. F. (1964). „Timbre cues and the identification of musical instruments“. *Journal of the Acoustical Society of America*, 36. S. 2021-2026.

Sandell, Gregory J. (1995). „Roles for spectral centroid and other factors in determining ‘blended’ instrument pairings in orchestration“. *Music Perception*, 13. S. 209-246.

Shepard, Roger. N. (1982). „Structural representations of musical pitch“. *The Psychology of Music*. Hrsg.: Deutsch, Diana. New York: Academic Press. S. 353-390.

Singh, Punita G. und Hirsh, Ira J. (1992). „Influence of spectral locus and F0 changes on the pitch and timbre of complex tones“. *Journal of the Acoustical Society of America*, 92. S. 2650-2661.

Sixtl, Friedrich. (1982). *Meßmethoden der Psychologie: Theoretische Grundlagen und Probleme*. 2. Auflage. Weinheim/Basel: Beltz Verlag.

Snyder, Bob. (2000). *Music and Memory: An Introduction*. Cambridge, Massachusetts: MIT Press.

Strong, William und Clark, Melville. (1967). „Synthesis of Wind-Instrument Tones“. *Journal of the Acoustical Society of America*, 41. S. 39-52.

von Hornbostel, Erich M. (1926). „Psychologie der Gehörerscheinungen“. *Handbuch der normalen und pathologischen Physiologie*, Bd. 11. Hrsg.: Bethe, A. S. 701-730.

Zwicker, Eberhard und Fastl, Hugo. (1999). *Psychoacoustics: Facts and Models*. 2. Auflage. Berlin: Springer Verlag.

# ANHANG A

## Tabelle mit den Einzelbewertungen der VP

Im folgenden werden aus dem Versuch die einzelnen Bewertungen von 21 VP über 66 Stimuli-Paare dargestellt. Die zwei Spalten ganz rechts geben die Mittelwerte bzw. Standardabweichungen über 20 VP (ohne VP *4f*) wieder. In der Systematik der VP-Namen kennzeichnen die Nummern 1 bis 4 hinsichtlich der Randomisierung im Versuch unterschiedliche Reihenfolgen von Paaren.

| VP | | 1a | 1b | 1c | 1d | 1e | 2a | 2b | 2c | 2d | 2e | 3a | 3b | 3c | 3d | 3e | 4a | 4b | 4c | 4d | 4e | 4f | $m(p_{ij})$ | $s(p_{ij})$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Alter | | 28 | 29 | 23 | 28 | 22 | 25 | 27 | 27 | 23 | 30 | 19 | 24 | 25 | 17 | 23 | 26 | 26 | 22 | 26 | 24 | 20 | | |
| Geschlecht | | m | w | w | m | w | m | m | m | w | m | w | m | m | w | w | w | m | w | w | w | m | | |
| Hörbeeinträchtigung | | n | n | n | n | n | n | j | n | n | n | n | n | n | n | n | n | n | n | n | n | n | | |
| Beispielpaar X | | 3 | 5 | 6 | 5 | 4 | 5 | 3 | 3 | 3 | 2 | 5 | 4 | 4 | 3 | 4 | 4 | 6 | - | 4 | 5 | 5 | 4.1053 | 1.1002 |
| Beispielpaar Y | | 4 | 4 | 4 | 4 | 3 | 3 | 3 | 5 | 2 | 4 | 5 | 2 | 3 | 4 | 5 | - | 6 | - | 5 | 1 | 6 | 3.7222 | 1.2744 |
| Proximität $p_{ij}$ | Paar | | | | | | | | | | | | | | | | | | | | | | | |
| $p_{2\|1}$ | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 4 | 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 3 | 1 | 2 | 1.4500 | 0.8870 |
| $p_{3\|1}$ | 2 | 2 | 1 | 3 | 1 | 1 | 2 | 1 | 2 | 3 | 3 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 4 | 1 | 3 | 1.7500 | 0.9105 |
| $p_{3\|2}$ | 3 | 2 | 3 | 3 | 2 | 1 | 2 | 1 | 4 | 1 | 3 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 4 | 1 | 2 | 1.9500 | 0.9987 |
| $p_{4\|1}$ | 4 | 2 | 2 | 4 | 4 | 4 | 5 | 3 | 4 | 3 | 1 | 4 | 4 | 5 | 5 | 3 | 3 | 5 | 5 | 5 | 5 | 6 | 3.8000 | 1.1965 |
| $p_{4\|2}$ | 5 | 3 | 5 | 5 | 5 | 5 | 5 | 3 | 6 | 6 | 2 | 3 | 3 | 6 | 6 | 4 | 3 | 4 | 5 | 4 | 4 | 6 | 4.3500 | 1.2258 |
| $p_{4\|3}$ | 6 | 3 | 5 | 2 | 4 | 5 | 6 | 2 | 5 | 5 | 3 | 5 | 2 | 4 | 4 | 3 | 6 | 5 | 5 | 5 | 6 | 6 | 4.2500 | 1.3328 |
| $p_{5\|1}$ | 7 | 4 | 2 | 5 | 3 | 4 | 4 | 3 | 3 | 5 | 2 | 5 | 5 | 6 | 6 | 6 | 3 | 4 | 4 | 3 | 5 | 6 | 4.1000 | 1.2524 |
| $p_{5\|2}$ | 8 | 3 | 4 | 5 | 4 | 4 | 3 | 3 | 5 | 5 | 4 | 3 | 4 | 6 | 6 | 5 | 4 | 5 | 5 | 5 | 5 | 6 | 4.4000 | 0.9403 |
| $p_{5\|3}$ | 9 | 3 | 5 | 6 | 4 | 5 | 6 | 2 | 6 | 5 | 4 | 5 | 5 | 5 | 6 | 2 | 4 | 5 | 5 | 4 | 5 | 6 | 4.6000 | 1.1877 |
| $p_{5\|4}$ | 10 | 1 | 1 | 1 | 1 | 3 | 1 | 2 | 4 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1.5500 | 0.8256 |
| $p_{6\|1}$ | 11 | 4 | 4 | 5 | 4 | 4 | 5 | 2 | 4 | 6 | 5 | 5 | 5 | 6 | 4 | 3 | 5 | 5 | 5 | 6 | 6 | 6 | 4.6500 | 1.0400 |
| $p_{6\|2}$ | 12 | 4 | 4 | 6 | 5 | 6 | 6 | 4 | 6 | 6 | 4 | 4 | 3 | 4 | 5 | 2 | 5 | 4 | 5 | 5 | 6 | 6 | 4.7000 | 1.1286 |
| $p_{6\|3}$ | 13 | 3 | 5 | 6 | 4 | 6 | 5 | 3 | 5 | 6 | 3 | 3 | 4 | 5 | 3 | 4 | 5 | 5 | 5 | 5 | 6 | 6 | 4.5500 | 1.0990 |
| $p_{6\|4}$ | 14 | 2 | 2 | 3 | 2 | 2 | 3 | 2 | 2 | 1 | 4 | 5 | 1 | 3 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 2 | 2.1000 | 1.0712 |
| $p_{6\|5}$ | 15 | 2 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 3 | 3 | 2 | 1 | 2 | 1 | 1 | 3 | 3 | 1 | 5 | 1 | 4 | 1.9500 | 1.0501 |
| $p_{7\|1}$ | 16 | 4 | 4 | 4 | 5 | 4 | 4 | 6 | 6 | 6 | 4 | 5 | 4 | 3 | 4 | 3 | 5 | 4 | 5 | 4 | 6 | 5 | 4.5000 | 0.9459 |
| $p_{7\|2}$ | 17 | 4 | 5 | 4 | 4 | 5 | 5 | 5 | 6 | 6 | 4 | 4 | 4 | 6 | 4 | 4 | 5 | 5 | 6 | 3 | 5 | 6 | 4.7000 | 0.8645 |
| $p_{7\|3}$ | 18 | 5 | 6 | 5 | 5 | 5 | 6 | 6 | 6 | 3 | 4 | 5 | 4 | 6 | 6 | 6 | 4 | 5 | 5 | 3 | 6 | 6 | 5.0500 | 0.9987 |
| $p_{7\|4}$ | 19 | 4 | 4 | 5 | 5 | 5 | 2 | 3 | 5 | 4 | 3 | 3 | 3 | 3 | 5 | 4 | 3 | 5 | 4 | 1 | 3 | 6 | 3.7000 | 1.1286 |
| $p_{7\|5}$ | 20 | 3 | 4 | 5 | 5 | 4 | 3 | 5 | 5 | 2 | 4 | 3 | 3 | 5 | 4 | 1 | 4 | 5 | 3 | 2 | 2 | 6 | 3.6000 | 1.2312 |
| $p_{7\|6}$ | 21 | 2 | 3 | 6 | 4 | 3 | 5 | 4 | 5 | 2 | 4 | 5 | 3 | 6 | 5 | 2 | 3 | 6 | 3 | 3 | 3 | 6 | 3.8500 | 1.3485 |
| $p_{8\|1}$ | 22 | 3 | 5 | 5 | 6 | 5 | 5 | 5 | 5 | 3 | 5 | 4 | 4 | 6 | 6 | 4 | 6 | 6 | 6 | 5 | 5 | 6 | 4.9500 | 0.9445 |
| $p_{8\|2}$ | 23 | 4 | 4 | 4 | 5 | 4 | 3 | 6 | 2 | 4 | 2 | 4 | 5 | 5 | 4 | 1 | 5 | 3 | 5 | 4 | 6 | 6 | 4.0000 | 1.2978 |
| $p_{8\|3}$ | 24 | 5 | 6 | 6 | 6 | 5 | 5 | 6 | 4 | 5 | 3 | 6 | 4 | 5 | 5 | 6 | 6 | 5 | 6 | 4 | 5 | 6 | 5.1500 | 0.8751 |
| $p_{8\|4}$ | 25 | 2 | 3 | 3 | 4 | 3 | 3 | 4 | 2 | 4 | 4 | 3 | 4 | 4 | 5 | 2 | 3 | 5 | 3 | 2 | 2 | 5 | 3.2500 | 0.9665 |
| $p_{8\|5}$ | 26 | 4 | 4 | 4 | 5 | 6 | 2 | 4 | 4 | 5 | 3 | 3 | 5 | 4 | 4 | 5 | 4 | 6 | 3 | 4 | 2 | 6 | 4.0500 | 1.0990 |
| $p_{8\|6}$ | 27 | 2 | 4 | 5 | 4 | 4 | 2 | 4 | 3 | 3 | 5 | 5 | 4 | 5 | 5 | 3 | 3 | 5 | 3 | 3 | 5 | 5 | 3.8500 | 1.0400 |
| $p_{8\|7}$ | 28 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 3 | 1.3000 | 0.4702 |
| $p_{9\|1}$ | 29 | 5 | 5 | 6 | 5 | 3 | 5 | 6 | 5 | 3 | 4 | 6 | 3 | 6 | 5 | 4 | 6 | 5 | 5 | 3 | 6 | 6 | 4.8000 | 1.1050 |
| $p_{9\|2}$ | 30 | 5 | 5 | 6 | 5 | 3 | 5 | 5 | 4 | 1 | 2 | 6 | 5 | 6 | 5 | 2 | 4 | 5 | 5 | 4 | 4 | 6 | 4.3500 | 1.3870 |
| $p_{9\|3}$ | 31 | 4 | 4 | 5 | 5 | 4 | 4 | 4 | 5 | 5 | 3 | 6 | 6 | 4 | 4 | 4 | 6 | 4 | 5 | 4 | 5 | 6 | 4.5500 | 0.8256 |
| $p_{9\|4}$ | 32 | 3 | 5 | 6 | 5 | 3 | 3 | 6 | 4 | 3 | 4 | 5 | 3 | 5 | 4 | 3 | 4 | 5 | 3 | 4 | 2 | 6 | 4.0000 | 1.1239 |
| $p_{9\|5}$ | 33 | 3 | 5 | 5 | 5 | 4 | 2 | 5 | 6 | 5 | 3 | 4 | 4 | 5 | 5 | 4 | 5 | 6 | 3 | 6 | 2 | 6 | 4.3500 | 1.2258 |

| VP | | 1a | 1b | 1c | 1d | 1e | 2a | 2b | 2c | 2d | 2e | 3a | 3b | 3c | 3d | 3e | 4a | 4b | 4c | 4d | 4e | ~~4f~~ | $m(p_{ij})$ | $s(p_{ij})$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Proximität $p_{ij}$** | **Paar** | | | | | | | | | | | | | | | | | | | | | | | |
| $p_{9\|6}$ | 34 | 3 | 5 | 5 | 5 | 5 | 3 | 5 | 6 | 4 | 2 | 5 | 5 | 3 | 4 | 5 | 2 | 5 | 3 | 4 | 4 | ~~5~~ | 4.1500 | 1.1367 |
| $p_{9\|7}$ | 35 | 3 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 4 | 6 | 1 | 4 | 1 | 2 | 1 | 3 | 1 | 1 | 4 | ~~4~~ | 2.3000 | 1.3416 |
| $p_{9\|8}$ | 36 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 4 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 3 | 2 | ~~2~~ | 1.7000 | 0.8013 |
| $p_{10\|1}$ | 37 | 4 | 4 | 3 | 4 | 5 | 5 | 6 | 4 | 6 | 5 | 5 | 5 | 5 | 4 | 2 | 6 | 5 | 6 | 4 | 6 | ~~6~~ | 4.7000 | 1.0809 |
| $p_{10\|2}$ | 38 | 4 | 6 | 5 | 5 | 4 | 5 | 6 | 6 | 5 | 4 | 4 | 5 | 6 | 5 | 4 | 3 | 5 | 6 | 4 | 6 | ~~6~~ | 4.9000 | 0.9119 |
| $p_{10\|3}$ | 39 | 4 | 4 | 4 | 4 | 3 | 6 | 5 | 6 | 4 | 3 | 6 | 5 | 6 | 4 | 6 | 3 | 6 | 6 | 2 | 5 | ~~6~~ | 4.6000 | 1.2732 |
| $p_{10\|4}$ | 40 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 6 | 4 | 4 | 3 | 5 | 2 | 2 | 2 | 5 | 5 | 4 | 4 | 4 | ~~6~~ | 3.6000 | 1.0954 |
| $p_{10\|5}$ | 41 | 2 | 3 | 5 | 3 | 4 | 2 | 3 | 4 | 2 | 3 | 3 | 3 | 5 | 2 | 6 | 3 | 5 | 4 | 3 | 3 | ~~5~~ | 3.4000 | 1.1425 |
| $p_{10\|6}$ | 42 | 3 | 2 | 5 | 4 | 2 | 5 | 4 | 6 | 2 | 5 | 6 | 5 | 2 | 2 | 2 | 5 | 5 | 4 | 4 | 2 | ~~6~~ | 3.7500 | 1.4824 |
| $p_{10\|7}$ | 43 | 3 | 3 | 1 | 3 | 4 | 3 | 2 | 3 | 4 | 2 | 4 | 4 | 4 | 3 | 1 | 4 | 3 | 2 | 4 | 3 | ~~2~~ | 3.0000 | 0.9733 |
| $p_{10\|8}$ | 44 | 1 | 3 | 3 | 4 | 2 | 2 | 2 | 4 | 3 | 4 | 2 | 4 | 2 | 3 | 5 | 4 | 4 | 2 | 1 | 2 | ~~4~~ | 2.8500 | 1.1367 |
| $p_{10\|9}$ | 45 | 4 | 2 | 4 | 5 | 3 | 3 | 3 | 6 | 2 | 3 | 5 | 3 | 5 | 3 | 5 | 2 | 5 | 2 | 1 | 4 | ~~6~~ | 3.5000 | 1.3572 |
| $p_{11\|1}$ | 46 | 4 | 4 | 5 | 5 | 5 | 4 | 6 | 6 | 6 | 4 | 6 | 5 | 4 | 5 | 4 | 4 | 5 | 5 | 4 | 4 | ~~6~~ | 4.7500 | 0.7864 |
| $p_{11\|2}$ | 47 | 3 | 4 | 3 | 4 | 4 | 4 | 5 | 6 | 5 | 4 | 6 | 3 | 6 | 6 | 5 | 5 | 5 | 6 | 4 | 5 | ~~6~~ | 4.6500 | 1.0400 |
| $p_{11\|3}$ | 48 | 4 | 6 | 6 | 5 | 5 | 6 | 5 | 5 | 5 | 4 | 6 | 4 | 6 | 4 | 5 | 5 | 5 | 5 | 4 | 4 | ~~6~~ | 4.9500 | 0.7592 |
| $p_{11\|4}$ | 49 | 2 | 4 | 4 | 3 | 4 | 3 | 5 | 4 | 4 | 3 | 4 | 3 | 3 | 2 | 2 | 3 | 5 | 4 | 4 | 3 | ~~6~~ | 3.4500 | 0.8870 |
| $p_{11\|5}$ | 50 | 2 | 4 | 6 | 4 | 5 | 2 | 5 | 3 | 2 | 2 | 4 | 5 | 3 | 3 | 2 | 4 | 4 | 4 | 5 | 2 | ~~6~~ | 3.5500 | 1.2763 |
| $p_{11\|6}$ | 51 | 3 | 4 | 6 | 5 | 4 | 3 | 5 | 5 | 4 | 5 | 4 | 2 | 4 | 3 | 2 | 5 | 6 | 4 | 5 | 3 | ~~6~~ | 4.1000 | 1.1653 |
| $p_{11\|7}$ | 52 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 3 | 1 | 4 | 1 | 3 | 2 | 2 | 3 | 2 | 2 | 2 | 2 | 1 | ~~1~~ | 1.9000 | 0.8522 |
| $p_{11\|8}$ | 53 | 2 | 2 | 1 | 3 | 2 | 1 | 2 | 2 | 2 | 1 | 3 | 4 | 2 | 2 | 2 | 2 | 3 | 2 | 1 | 1 | ~~3~~ | 2.0000 | 0.7947 |
| $p_{11\|9}$ | 54 | 3 | 3 | 3 | 4 | 2 | 3 | 1 | 4 | 2 | 3 | 1 | 4 | 2 | 1 | 3 | 3 | 3 | 1 | 6 | 1 | ~~3~~ | 2.6500 | 1.3089 |
| $p_{11\|10}$ | 55 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 4 | 3 | 5 | 2 | 1 | 3 | 1 | 1 | 1 | 3 | 2 | 1 | 3 | ~~2~~ | 2.1000 | 1.1192 |
| $p_{12\|1}$ | 56 | 5 | 5 | 6 | 6 | 5 | 3 | 5 | 5 | 4 | 5 | 6 | 5 | 6 | 6 | 4 | 5 | 6 | 6 | 2 | 6 | ~~6~~ | 5.0500 | 1.0990 |
| $p_{12\|2}$ | 57 | 4 | 5 | 6 | 5 | 5 | 6 | 5 | 6 | 4 | 4 | 6 | 3 | 6 | 6 | 4 | 6 | 4 | 6 | 5 | 6 | ~~6~~ | 5.1000 | 0.9679 |
| $p_{12\|3}$ | 58 | 4 | 4 | 4 | 5 | 4 | 5 | 3 | 5 | 3 | 4 | 4 | 4 | 6 | 5 | 6 | 5 | 5 | 6 | 5 | 3 | ~~6~~ | 4.5000 | 0.9459 |
| $p_{12\|4}$ | 59 | 3 | 3 | 5 | 5 | 5 | 2 | 5 | 6 | 2 | 4 | 4 | 2 | 4 | 3 | 1 | 4 | 5 | 4 | 3 | 3 | ~~6~~ | 3.6500 | 1.3089 |
| $p_{12\|5}$ | 60 | 3 | 4 | 5 | 5 | 4 | 4 | 5 | 4 | 1 | 3 | 6 | 3 | 3 | 4 | 1 | 3 | 6 | 4 | 2 | 3 | ~~6~~ | 3.6500 | 1.3870 |
| $p_{12\|6}$ | 61 | 3 | 3 | 4 | 4 | 3 | 3 | 5 | 5 | 4 | 3 | 3 | 3 | 3 | 4 | 1 | 2 | 5 | 4 | 4 | 4 | ~~5~~ | 3.5000 | 1.0000 |
| $p_{12\|7}$ | 62 | 2 | 2 | 2 | 2 | 2 | 5 | 1 | 2 | 1 | 4 | 2 | 2 | 2 | 3 | 2 | 1 | 3 | 1 | 1 | 4 | ~~4~~ | 2.2000 | 1.1050 |
| $p_{12\|8}$ | 63 | 3 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 3 | 4 | 3 | 1 | 2 | 2 | 2 | 1 | 3 | 2 | 1 | 1 | ~~2~~ | 1.9000 | 0.9119 |
| $p_{12\|9}$ | 64 | 2 | 1 | 1 | 3 | 2 | 1 | 3 | 3 | 4 | 1 | 2 | 2 | 3 | 2 | 4 | 2 | 3 | 2 | 1 | 4 | ~~2~~ | 2.3000 | 1.0311 |
| $p_{12\|10}$ | 65 | 4 | 2 | 5 | 4 | 3 | 2 | 3 | 4 | 2 | 5 | 5 | 3 | 5 | 3 | 3 | 3 | 4 | 1 | 1 | 4 | ~~6~~ | 3.3000 | 1.2607 |
| $p_{12\|11}$ | 66 | 2 | 1 | 2 | 2 | 3 | 1 | 3 | 3 | 2 | 5 | 2 | 1 | 3 | 2 | 1 | 1 | 3 | 2 | 2 | 2 | ~~2~~ | 2.1500 | 0.9881 |

**Anmerkungen**

$m(p_{ij})$: arithmetischer Mittelwert über alle VP (ohne VP *4f*)

$s(p_{ij})$: Standardabweichung über alle VP (ohne VP *4f*)

| | | |
|---|---|---|
| Maximum | 5.1500 | 1.4824 |
| Minimum | 1.3000 | 0.4702 |
| Mittel | 3.6061 | 1.0720 |