

Der Klang der Marken – Untersuchungen zu branchentypischen Eigenschaften von Audiologos

Christoph Anzenbacher, Isabella Czedik-Eysenberg,
Christoph Reuter & Michael Oehler

Zusammenfassung

Auf der Grundlage einer (psycho-)akustischen und musikalischen Analyse werden im vorliegenden Beitrag die typischen klanglichen Eigenschaften von 200 Audiologos ermittelt. Mithilfe der via MATLAB/MIRtoolbox ermittelten Ergebnisse lassen sich hierbei die Audiologos in vielen Fällen allein schon aufgrund ihrer klanglichen Eigenschaften bestimmten Branchen oder Industriezweigen zuordnen. Besonders am spektralen Schwerpunkt (Spectral Centroid) und am Einschwingvorgang lassen sich direkte Bezüge zwischen klanglicher Ausgestaltung und Branchenzugehörigkeit feststellen. Mit anderen Worten: Vom Klang eines Audiologos lässt sich in vielen Fällen direkt auf seine Branchenzugehörigkeit schließen. Auf der Grundlage der vorliegenden Ergebnisse lassen sich somit direkte Vorgaben für ein branchencharakteristisches Sounddesign von neuen Audiologos erstellen.

Abstract

In the present study typical sound properties of 200 audio logos are determined by means of (psycho-)acoustical and musical analyses. While using the results of an audio feature analysis via MATLAB/MIRtoolbox most audio logos can be assigned to a specific sector or branch of industry. The parameters spectral centroid and first attack time are especially suited to determine correlations between sound properties and branches of industry. In other words: In many cases the sound of an audio logo can be connected directly to its industry affiliation. On that base the results of the study can be used in the field of audio branding to design characteristic and customized audio logos for selected branches of industry.

1 Hintergrund

Audiologos versprechen als akustische Pendants zum visuellen Markenlogo häufig einen hohen verkaufssteigernden Wiedererkennungswert für Produkte,

Dienstleistungen und Organisationen aller Art. Hauptsächlich positioniert am Beginn oder Ende von Werbespots soll das Audiologo als aufmerksamkeitssteigerndes Element durch häufige Wiederholungen auf die beworbene Markenidentität hinweisen und diese bekannt machen. Neben Einsatz und Wirkung von Musik in der Werbung (u. a. Allan, 2007; Kellaris & Kent, 1991; Shen & Chen, 2006) wurden in den letzten Jahren die meist nur wenige Sekunden kurzen akustischen (Marken-)Signale in einer Reihe von Studien hinsichtlich ihrer musikalischen Passung, melodischen Eingängigkeit, ihrer crossmodalen Wirkungsweise, ihres aufmerksamkeitssteigernden Effekts und ähnlicher Faktoren untersucht (Anzenbacher, 2012; Anzenbacher et al., 2013; Bronner & Hirt, 2009; North et al., 2004). Dabei konnte die Frage nach der branchenspezifischen Differenzierung aufgrund akustischer Eigenschaften dieser akustischen Werbeträger bisher noch nicht beantwortet werden.

Deswegen wird in der vorliegenden Studie den Fragen zur bewussten Auswahl, zum gezielten Einsatz oder möglichen Wirkungsebenen akustischer Merkmale als wesentliches Ausgangsmaterial branchenspezifischer Markengestaltung auf den Grund gegangen.

Eine der wesentlichsten Anforderungen an ein Audiologo ist die vom Wettbewerb geforderte Unterscheidbarkeit von anderen Audiologos. Wesentlich dafür und damit für den Werbeerfolg ist das aufmerksamkeitserregende Potenzial, das die Marke oft durch die Andersartigkeit im Vergleich der Mitbewerber erreichen kann. Auf visueller Wahrnehmungsebene grenzen sich Logos durch ihre Gestaltung voneinander ab und kommunizieren über strategisch konkret eingesetzte semantische Zusammenhänge, wie etwa die Farbsymbolik, Formen oder einer indexikalischen Gestaltung, die Ähnlichkeiten mit der bezeichneten Leistung aufweist. Die Marke Telekom erscheint in auffälligem Magenta, Nivea in unverkennbarem vertrauensvollem Blau, Allianz lässt mit einem stilisierten Adler Assoziationen zur starken Firmentradition zu, UBS bekräftigt mit den drei sternförmig angeordneten Schlüsseln die einstigen Werte Sicherheit, Vertrauen und Verschwiegenheit.

Nun sind der Gestaltung von Audiologos aufgrund der linearen auditiven Reizverarbeitung und dem Einsatz hauptsächlich in zeitlich limitierten Werbespots naturgemäß strukturelle Grenzen gesetzt. Sind es für gewöhnlich charakteristische Melodielinien und harmonische Fortschreitungen, die dem Hörer als Unterscheidungsmerkmal dienen, rückt bei Audiologos speziell auch die klangliche Textur (klangliche Umsetzung, z. B. Instrumentierung) in den Vordergrund. Innerhalb nur weniger Sekunden soll beim Hörer durch konkrete Assoziationen ein möglichst geschlossenes Markenbild entstehen, wie es auch der Anspruch der visuellen Inszenierung ist. Das bedeutet für den Komponisten, auf einfach verständliche Formen (z. B. „hell“ vs. „dunkel“) zurückzugreifen, die analog zur visuellen Gestaltung u. a. den Bedeutungstransfer durch semantische Inhalte unterstützen können und so Facetten der Markenpersönlichkeit bzw. gewünschte Eigenschaften der Produkte hervorheben können. Da innerhalb von Branchen und Produktgruppen gleiche Eigenschaften existieren, scheinen Fragen nach typischen Gemeinsamkeiten auch in der Textur von Audiologos durchaus interessant:

- Gibt es also bestimmte Klangeigenschaften, die besonders typisch für Audiologos sind?
- Lassen sich branchentypische klangliche Merkmale ermitteln?
- Welche Melodieverläufe sind besonders charakteristisch und wie werden die Stimme und der Text einbezogen?

Eine umfangreiche vergleichende (psycho-)akustische und musikalische Analyse der klanglichen Merkmale von 200 Audiologos aus 10 verschiedenen Branchen soll über die charakteristischen klanglichen Eigenschaften von Audiologos und deren branchenspezifischen Ausformung Aufschluss verleihen¹. Die Kenntnis möglicher branchenspezifischer Merkmale kann u.U. hilfreich für die strategische Klangentwicklung und die kreative Arbeit von Komponisten oder Sound Designern sein.

2 Methode

Der Studie voraus ging eine Bestandsaufnahme der am internationalen Markt existierenden Audiologos (mit einem Schwerpunkt auf europäische Audiologos) auf der Grundlage ihres Erscheinens in TV- und Radioaufnahmen, ihrer Behandlung in wissenschaftlicher Literatur (u. a. Bronner & Hirt, 2009; Jackson, 2004; Steiner, 2014) und internationalen Markenämtern. Alle Beispiele stammen – sofern sie nicht isoliert zur Verfügung standen – von Werbespots (Video und Audio) und wurden extrahiert und normalisiert (*.wav, 44.1 kHz, 16-bit, Stereo).

2.1 Auswahl der Stimuli

Der Fokus dieser Studie ist klar auf das Audiologo (als kürzester akustischer Repräsentant einer Marke) gerichtet. Da häufig aber nur schwer zwischen verschiedenen Formen des Audio-Brandings unterschieden werden kann und die Grenzen zwischen Audiologo und verwandten akustischen Werbeträgern oftmals sehr fließend sind, erfolgte die Auswahl und Definition der eindeutig als Audiologo zu bezeichnenden Stimuli aufgrund der folgenden Kriterien:

- Ein Audiologo lässt sich eindeutig einer Marke, einer Dienstleistung oder einer Organisation zuordnen.
- Die vom Audiologo repräsentierte Marke lässt sich eindeutig einem wirtschaftlichen Sektor („industries“) zuordnen.
- Ein Audiologo hat eine Länge von maximal 20 Sekunden.

1 Man mag an dieser Stelle einwenden, dass die getroffene Auswahl von 200 Audiologos aus 10 verschiedenen Branchen zwar in vielen aber vielleicht nicht in allen der in Folge behandelten Fälle eine genügend große Stichprobengröße aufweist. Da es sich hierbei um die jeweils bekanntesten Audiologos pro Branche handelt (und es auch je nach Branche gar nicht so viele Audiologos gibt) lässt sich auch bei diesen vereinzelt Fällen von einer gewissen Repräsentativität der Ergebnisse ausgehen.

Nach diesen Kriterien wurde eine Auswahl von 200 Audiologos zusammengestellt (mit einem Fokus auf vorhandene Audiologos der weltweit wertvollsten Marken², die sich vom Kerngeschäft des Unternehmens/der Marke her mit zehn verschiedenen Branchenbezeichnungen (z. B. Automobil, Telekommunikation) kategorisieren ließen. Zusätzlich zu dieser Grundkategorisierung wurde die internationale Standardklassifikation der Wirtschaftszweige (United Nations Statistics Division, 2014) herangezogen, in der jede Branche einer von 21 Hauptgruppen zugeordnet wird, die wiederum in weitere Subkategorien aufgeteilt sind. Die Entscheidung für diese Klassifikation beruht darauf, dass dieser Standard weltweit eingesetzt wird und eine genaue und standardisierte Branchenunterscheidung erlaubt. So lassen sich die Marken nicht nur einer übergeordneten Hauptgruppe (z. B. Verarbeitendes Gewerbe [Manufacturing]) zuweisen, sondern auch Sparten (z. B. Nahrungsmittelproduktion [Manufacture of food products]) und Untergruppen (z. B. Manufacture of bakery products), wodurch dieses System als sehr flexibel und stabil betrachtet werden kann. Alternative Klasseneinteilungen zeichnen sich entweder durch ein äußerst differenziertes System aus (z. B. Nizza Klassifikation für Waren und Dienstleistungen³ mit 45 Klassen) oder können zwar als etabliert, aber trotz der fehlenden Dokumentation als wenig konsistent bewertet werden (z. B. Klassifikation internationaler Organe wie Interbrand, Fortune) und wurden deswegen verworfen. So folgt die Aufgliederung in Branchenklassen via United Nations Statistics Division durch ihre hierarchische Struktur einem systematischen Ansatz. Jedoch unterliegt sie auch der subjektiven Zuordnung einer Marke zu einer Haupt- und Untergruppe und damit der zugeschriebenen Herkunftsbranche. In Einzelfällen und gerade bei Dach- und Unternehmensmarken (z. B. *Siemens*, *Sony*) und im Bereich Retail (G) wären einerseits (mehrfache) Zuordnungen zu alternativen Kategorien möglich, andererseits eine Zusammenfassung (z. B. Luft- und Bahnlinien [Transport, H] unter Travel [N]), wobei in dieser Studie die Zuordnung immer für nur eine Hauptgruppe gemäß der Klassifikationsbeschreibung erfolgt. Die Bindung an eine Klassifikation kann prinzipiell zu Varianzen oder widersprüchlichen Effekten führen. Einschränkungen in der Reliabilität sind auch durch Revisionen der Klassifizierungssysteme gegeben, die im regelmäßigen Turnus veröffentlicht werden und in denen sich ggf. die Markenzuordnung innerhalb einer Untergruppe abwandeln können. Der Einbezug solcher Änderungen ist natürlich unentbehrlich, was wiederum für die Verwendung gröberer und eher praxisnaher dafür aber konsistenterer Branchenkategorien spricht. Die Eignung des richtigen Verfahrens sollte darum im Einzelfall diskutiert werden. Aus der Auswahl der verschiedenen Klassifizierungsansätze kristallisierte sich aus den oben genannten Gründen die Klassifikation über die United Nations Statistics Division als die geeignetste heraus. Hinsichtlich der Historizität der Audiologos sind zwar Hinweise bei den in den Patentämtern eingetragenen Hörmarken zu finden. Zum

2 Siehe z. B. Methodologie und Ranglisten aus der Marktforschung (z. B. BrandZ von MillwardBrown) oder der Markenberatung (z. B. Best Brands von Interbrand).

3 Zugriff unter https://www.dpma.de/docs/service/klassifikationen/nizza/nizza_11-2017_einleitung_klassentitel_klasseneinteilung.pdf

einen unterliegen die Zeitangaben aber nur dem aktiven Anmeldeprozess (so, dass der reale Einsatz der Audiologos anhand dieser Daten nur vermutet werden kann), zum anderen konnten die Informationen für die ausgewählte Population auch aus anderen Quellen nur lückenhaft dokumentiert werden. Die folgende Analyse kann somit nur unabhängig vom Zeitpunkt der Produktion bzw. Verwendung erfolgen.

2.2 Auswahl der Analyseparameter

Die als Audiosamples normalisiert vorliegenden Stimuli wurden (nach einer automatisierten Bestimmung der Länge sowie einer groben Segmentierung) auf eine Reihe von klanglichen Eigenschaften hin untersucht (vgl. Tab. 1). Die damit zusammenhängenden Analysen und Berechnungen wurden mithilfe der MIR-toolbox in MATLAB durchgeführt (Lartillot & Toivianen, 2007; Lartillot, Toivianen & Eerola, 2008).

Hier war ein wesentlicher Deskriptor der Spectral Centroid: Als gewichteter Mittelwert des Frequenzspektrums korreliert er stark mit der empfundenen „Hellichkeit“ eines Klanges und lässt sich daher als grundlegendes Merkmal zum Vergleich der Klangfarbe zweier Klänge heranziehen. Zusätzlich wurden spektrale Fluktuationen betrachtet. Der entsprechende Deskriptor (Spectral Flux) berechnet sich als euklidische Differenz des Spektrums jedes Frames zu dem des vorhergehenden. Hierbei kamen halbüberlappende Frames mit einer Länge von 50 ms zum Einsatz. Die resultierenden Differenzen wurden schließlich gemittelt, um ein quantitatives Maß für das insgesamt Auftreten von spektralen Schwankungen innerhalb eines Audiologos zu erhalten. Die Berechnung der Rauigkeit (Roughness) erfolgt – basierend auf einem Algorithmus von Sethares (1998) –, indem die Peaks des Spektrums bestimmt und anschließend paarweise in jeder möglichen Kombination auf Dissonanz untersucht werden: Ob zwei Sinusoide gemeinsam eine rauigkeitsverursachende Amplitudenmodulation aufweisen, wird dabei nach ihrem Frequenzverhältnis zueinander gemäß der Beschreibung von Plomp und Levelt (1965) bestimmt. Der Durchschnitt aller so ermittelten Einzeldissonanzen dient anschließend zur Beschreibung der sensorischen Dissonanz des Gesamtklantes (Lartillot, 2013, S. 127).

Als weitere Eigenschaft wurde die Inharmonizität jedes Klangbeispiels herangezogen. Deren Abschätzung geschieht, indem auf das Signal eine Filterfunktion angewandt wird, durch welche sich der Energieanteil jener Frequenzkomponenten messen lässt, die nicht im Bereich der Vielfachen der Grundfrequenz liegen (Lartillot, 2013, S. 137). Ebenfalls wurde der Niedrigenergieanteil (Low Energy Rate) erfasst; Dieser bezeichnet den Prozentsatz aller Frames, deren Energie unter dem Durchschnitt liegt (Tzanetakis, 2002). Dies dient dazu, das Verhältnis des Auftretens ruhigerer und energiereicherer Zeitfenster zu ermitteln. Wesentlich für die Analyse waren auch die Einschwingzeiten der verwendeten Klänge. Im Kontext dieser Arbeit wurde – hinsichtlich der unterschiedlichen Länge und Struktur der Audiologos – im Besonderen die Länge des ersten vor kommenden Einschwingvorgangs als vergleichbarer Parameter herangezogen.

Tab. 1:
Liste der untersuchten klanglichen Eigenschaften (timbre features)

| Timbre Feature | Kurzdefinition | Empfindung |
|-------------------|--|---|
| Spectral Centroid | Spektraler Schwerpunkt (nach Amplituden gewichteter spektraler Mittelwert der pro Frame im Spektrum vorhandenen Frequenzen) | Korreliert mit der Helligkeitsempfindung: je höher der Spectral Centroid, desto heller der Klang. |
| Spectral Flux | Spektrale Veränderung (Maß für die Stärke von klanglichen Schwankungen, berechnet als euklidische Distanz der Spektren aufeinanderfolgender Zeitfenster) | Je höher der Spektral Flux, desto bewegter der Klang (gut zur Erkennung von Toneinsätzen oder als Indikator für das Vorhandensein von Sprache). |
| Roughness | klangliche Rauigkeit (Maß für die Anwesenheit von Modulationen unter 300 Hz) | Je stärker die Modulationen, desto rauer der Klang. |
| Inharmonicity | Klangliche Inharmonizität (Maß für die Anwesenheit von nicht zur Obertonreihe passenden Teiltönen) | Je höher die Inharmonizität, desto unharmonischer, verzerrter der Klang. |
| Low Energy Rate | Niedrigenergieanteil (Maß für den Anteil von Frames mit einem unterdurchschnittlichen Energieinhalt) | Je höher die Low Energy Rate, desto eher weist der Klang einzelne Lautstärkemaxima unter vielen ruhigen Zeitfenstern auf. |
| First-AttackTime | Erste Einschwingzeit (Maß für die Zeit, in der die Amplitude des Klangs zum ersten Mal sein erstes Maximum erreicht) | Je kürzer die erste Einschwingzeit, desto abrupter der Einsatz. |
| Dynamic Range | Dynamische Bandbreite (Maß für dynamische Schwankungen, bzw. für die Stärke einer dynamischen Kompression, berechnet als Abstand zwischen Maximum und 0.1-Perzentil.) | Je geringer die dynamische Bandbreite, desto weniger kann zwischen laut und leise unterschieden werden. |
| Unangenehmheit | klangliche Unangenehmheit/Schrillheit (Maß für Ausprägtheit einer Tonhöhe verbunden mit der Anwesenheit von starken Amplituden bei 2–4 kHz) | Je ausgeprägter der Tonhöhereindruck und je stärkere Anteile zwischen 2–4 kHz, desto unangenehmer, schriller, quietschender der Klang. |

Hierfür wurde der zeitliche Abstand zwischen dem ersten lokalen Maximum und dem letzten davor auftretenden Minimum in der Hüllkurve des Signales bestimmt, mit dem Ziel, auch im Fall eines gemischten Klanges gewisse Information darüber zu erlangen, wie abrupt der Anfang eines Audiologos sich vollzieht. Weiterhin standen die dynamischen Schwankungen der vorliegenden Klangdateien im Mittelpunkt des Interesses, weswegen eine rechnerische Einschätzung der absoluten bzw. einer relativen dynamischen Bandbreite durchgeführt wurde: Nach einer frameweisen Bestimmung der rms-Werte wurde der Abstand des Energiemaximums zum 0.1-Quantil berechnet. Im Falle der relativen dynamischen Bandbreite wurde diese Differenz anschließend durch den Median der rms-Energie normalisiert.

Als zusätzlicher experimenteller Deskriptor wurde die physiologische „Unangenehmheit“ automatisiert für die vorliegenden Klänge eingeschätzt, wie sie von Reuter und Oehler in ihrer Studie zu Wandtafelkratzgeräuschen (Reuter & Oehler, 2011) untersucht wurde. Gemäß den klanglichen Faktoren, die sich für einen unangenehmen Eindruck als besonders ausschlaggebend erwiesen (ausgeprägte Tonhöhenanteile, insbesondere in Zusammenhang mit starken Frequenzanteilen zwischen 2000 und 4000 Hz) wurde ein entsprechender Deskriptor implementiert. Dieser setzt sich als Produkt des im genannten Frequenzbereich vorliegenden Energieanteils und der Stärke der harmonischen Tonhöhenanteile, welche als Gegenstück zu den inharmonischen Komponenten (siehe oben) abgeschätzt wurden, zusammen.

Hinzu kam die kategoriale Erfassung des Herkunftslands der jeweiligen Marke (Unternehmenssitz) und weiterer klanglicher/musikalischer Eigenschaften mittels Expertenanalyse. Zu den Eigenschaften wurde der strukturelle melodische Konturverlauf herangezogen, das Auftreten (occurrence) und die Beschaffenheit (manner) von Sing- und Sprechstimme(n) (dazu zählen auch Pfeifen, Flüstern und ähnliche Lauterzeugungen⁴). Die Erfassung und Verschlagwortung der am jeweiligen Audiologo beteiligten Musikinstrumente basierte auf schriftlichen Angaben z. B. aus Markeneintragungen – sofern diese zur Verfügung standen – und Erfassung durch Experten aufgrund analytischem Hören der Audiologos über Kopfhörer. Zur Bestimmung der Länge wurde ebenfalls die Anzahl der Events (klanglich trennbare Elemente in zeitlicher Reihenfolge) der verschiedenen Audiologos transkribiert.

Es wurden bewusst mehrere dieser strukturellen Merkmale ausgewählt, da sie auf eine einfache Weise durch vergleichende Betrachtung eine Beurteilung der klanglichen Unterscheidbarkeit von Audiologos (und von Marken) vermuten lassen.

Da eine Kernanforderung an die Gestalt eines Audiologo darin besteht, unmissverständlich wahrnehmbar und erinnerbar zu sein, sollte das Audiologo eine einfache Phrase bilden und die Grenzen der Kapazität des musikalischen Kurzzeitgedächtnisses nicht überschreiten (Snyder, 2000; Müllensiefen, 2004). Darum soll hier der Parameter *Kontur* exemplarisch hervorgehoben werden. Nach

4 Für einen Überblick zu Arten vokaler Inszenierung siehe Anzenbacher (2016, S. 74).

den Untersuchungen von Dowling (1978), Eiting (1984), Cutietta und Booth (1996) u. a. hat sich die melodische Kontur (gemeinsam mit der Tonhöhenwahrnehmung) als entscheidend für eine bewusste musikalische Informationsverarbeitung erwiesen. So gelten Phrasen mit symmetrischem Aufbau und geringer Richtungsänderung als besonders hilfreich für die musikalische Memorierung einfacher melodischer Gestalten. Da bereits beobachtet werden konnte, dass Audiologos mit ihrem Motivcharakter eine durchschnittliche Anzahl von fünf Tönen (bzw. musikalischen Events) aufweisen (Anzenbacher, 2012), lassen sich im Falle der Audiologos besonders gut die musikalischen Konturtypen nach David Huron auf die vorliegenden Klangbeispiele anwenden. Mit den „aggregate phrase contours“ (Huron, 1996) definierte Huron insgesamt neun Typen von melodischen Konturen. Fünf von diesen Konturtypen (*convex*⁵, *concave*, *ascending*, *descending*, *horizontal*) wurden als Kategorien für die Bestimmung der klanglich-melodischen Konturen in der vorliegenden Studie verwendet. Für konturlose Einzeltöne wurde die Einteilung *single event* vergeben (falls ein Audiologo nur aus einem musikalisch-akustischen Event bestand) und *wavelike* als vereinfachende Zusammenfassung der Bezeichnung von mehrfach auf- und absteigenden Tonfolgen (von Huron, 1996, S. 9ff., verwendete Konturtypen: *horizontal-descending*, *horizontal-ascending*, *ascending-horizontal* und *descending-horizontal*). Jedem Audiologo wurde nur ein Konturtyp zugeordnet, eine Kombination aus mehreren Konturtypen wird nicht berücksichtigt. Ein besonderer Vorteil dieser Konturtypen bestand darin, dass sie unabhängig von der rhythmischen Dimension, der Länge der einzelnen tonalen Events und des Ambitus⁶ des jeweiligen Audiologos eingesetzt werden konnten.

Obwohl auch die Rhythmik als eines der wichtigsten Merkmale in der auditiven Wahrnehmung von Audiologos hohe Relevanz hat, werden Deskriptoren für rhythmische Merkmale in dieser Studie vorerst bewusst ausgeklammert, sollen aber in kommenden Arbeiten auch hinsichtlich der Branchenunterschiede berücksichtigt werden. Aufgrund der Kürze der Beispiele ist es oftmals schwierig, ein Takmaß und damit einhergehend eine rhythmische Struktur zu bestimmen. Eine Möglichkeit für die Kodierung bietet sich allerdings in einer groben rhythmisch-semantischen Beschreibung und der Erfassung des stilistischenindrucks, d. h. ob eher eine binäre „rockige“, eine ternäre „swingende“ Stilistik oder andere Ausprägungen wie etwa ein Offbeat vorliegen.

3 Ergebnisse

3.1 Allgemeines

Die 200 verwendeten Audiologos entstammten aus zehn übergeordneten Industriezweigen mit Herkunftsländern aus Amerika, Europa und Asien, wobei der

⁵ Kategorie Convex (oder gewölbte Kontur) besitzt nicht zwingend den gleichen Anfangs- und Endton.

Schwerpunkt vor allem auf dem amerikanischen und europäischen Raum liegt. So stammen die Audiologos europäischer Herkunft aus Belgien, Dänemark, Deutschland, Finnland, Frankreich, Italien, Luxemburg, Niederlande, Österreich, Polen, Slowakei, Spanien, Schweden, Schweiz und England, während die amerikanischen Audiologos aus Brasilien, Kanada und den USA stammen sowie die asiatischen aus Indonesien, Japan und Südkorea (vgl. Tab. 2).

Tab. 2:
Verteilung der Audiologos innerhalb der Branchen und Kontinente

| Industry | America | Asia | Europe | Total |
|---------------------------|------------|-------------|-------------|--------------|
| C-10 (food) | 3 1.5 % | 1 .5 % | 16 8 % | 20 10 % |
| C-11 (beverage) | 1 .5 % | 0 0 % | 12 6 % | 13 6.5 % |
| C-26 (electronic) | 8 4 % | 13 6.5 % | 8 4 % | 29 14.5 % |
| C-29 (automotive) | 1 .5 % | 5 2.5 % | 15 7.5 % | 21 10.5 % |
| D (energy) | 3 1.5 % | 0 0 % | 14 7 % | 17 8.5 % |
| G (retail) | 2 1 % | 0 0 % | 24 12 % | 26 13 % |
| J-59 (motion pictures) | 12 6 % | 1 .5 % | 2 1 % | 15 7.5 % |
| J-60 (broadcasting) | 5 2.5 % | 0 0 % | 11 5.5 % | 16 8 % |
| J-61 (telecommunication) | 3 1.5 % | 0 0 % | 12 6 % | 15 7.5 % |
| K-64 (financial services) | 2 1 % | 0 0 % | 26 13 % | 28 14 % |
| Total | 40 20 % | 20 10 % | 14 70 % | 200 100 % |

3.2 Instrumentation

Allein von der Verwendung der Musikinstrumente lässt sich schon häufig auf die Branchenzugehörigkeit eines Audiologos schließen: So werden elektronische, synthetisch erzeugte Klänge mit Abstand am häufigsten eingesetzt. Den Synthesizer findet man erwartungsgemäß vor allem in technisch orientierten

Bereichen wie der Elektronikbranche (C-26) und dem Rundfunk (J-60) – hier auch oft in Kombination mit stark geräuschhaften Klängen, Streichern und Piano (z. B. *Intel*, *Samsung*, *Gigaset*, *Kabel Eins*, *Olympus*). In Branchen, in denen es um die Natürlichkeit der Produkte geht (z. B. Lebensmittel, C-10), findet man den Synthesizer dagegen nur vereinzelt. Die Lebensmittelbranche (C-10) ist dagegen besonders von (eher weiblichen) Gesangsstimmen sowie Schlag- und Saiteninstrumenten geprägt (z. B. *Activia*, *Exquisa*, *Yakult*). Auffallend ist, dass männliche Stimmen vor allem wiederum die Audiologos aus dem Einzelhandel (G) prägen (z. B. *Ebay*, *Edeka*, *Obi*, *Praktiker*, *Toys “R” Us*), während weibliche Stimmen eher in den Branchen der Ernährung (C-10) und Energie (D) zu hören sind (z. B. *Activia*, *RWE*, *DEA*) und in eher „männlich“ orientierten Branchen wie Elektronik (C-25) oder Getränke (C-11) nahezu verstummen. Klänge der E-Gitarre hört man besonders häufig in Audiologos der Automobilindustrie (C-29) und des Rundfunks (J-60) (z. B. *Ford*, *Kia*, *Astral Media*, *HLN*), während das eher akzentsetzende Glockenspiel (vielleicht als Reminiszenz zur Telefonklingel oder zum Signalton) im Telekommunikationsbereich (J-61) erklingt (z. B. *Drei*, *Mobilcom Debitel*). Die von ihrer Tradition her eher als hoheitsvoll, ernst und kühl (Bruner, 1990, S. 97) charakterisierten Blechblasinstrumente begegnen einem auch bei den Audiologos besonders im Bereich des Films (J-59) und des Rundfunks (J-60) (z. B. *CBS*, *Nickelodeon*) sowie bezeichnenderweise auch im Finanzdienstleistungssektor (K-64) (z. B. *DEVK*, *Loomis*). Letzterer beheimatet aber auch ohnehin die größte Instrumentenvielfalt (mit besonderer Bevorzugung des Klaviers), während im Bereich der Elektronik (C-25) charakteristischerweise zum größten Teil synthetische Klänge bzw. stark modifizierte Instrumentenklänge verwendet werden (z. B. *Intel*, *HP*, *Quimonda*, *Namco*) (vgl. Abb. 1).

3.3 Kontur

Die Kontur des Audiologos beschreibt den Melodieverlauf, ob dieser eher auf- oder absteigend bzw. konvex oder konkav ist oder wellenförmig oder gleichbleibend horizontal verläuft. Wie ein Chi-Quadrat-Test zeigt (χ^2 [45, $N=196$] = 69.22; $p=.012$; $\phi=0.594$), unterscheidet sich die Kontur der Audiologos signifikant in Abhängigkeit von der Branche, nicht jedoch in Abhängigkeit vom Kontinenten (χ^2 [10, $N=196$] = 13.524; $p=.196$; $\phi=0.263$) (vgl. Tab. 3).

So können Audiologos mit einer wellenförmigen melodischen Kontur (wave-like) besonders im Bereich der Finanzdienstleistungen (K-64), Elektronik (C-25), der Film- (J-59) und der Lebensmittelindustrie (C-10) gehört werden, während die aufsteigende Melodiekurve eindeutig die Energiebranche (D) beschreibt (z. B. *Entega*, *EnBW*, *Vattenfall*). Absteigende oder konkave Melodiekonturen finden sich vergleichsweise weniger; wenn, dann vor allem im Einzelhandel (G) (z. B. *Chicco*, *Reno*). Konvexe Melodiekonturen (die also erst aufwärts und dann wieder abwärts verlaufen) sind besonders häufig in der Telekommunikationsbranche (J-61) (z. B. *Bob*, *Drei*, *Telekom*, *Red Bull Mobile*) vertreten sowie im Rundfunk (J-60), in der Film- (J-59), Getränke- (C-11) und Automobilindustrie (C-29). In letzterer begegnet man auch vergleichsweise oft einem horizontalen

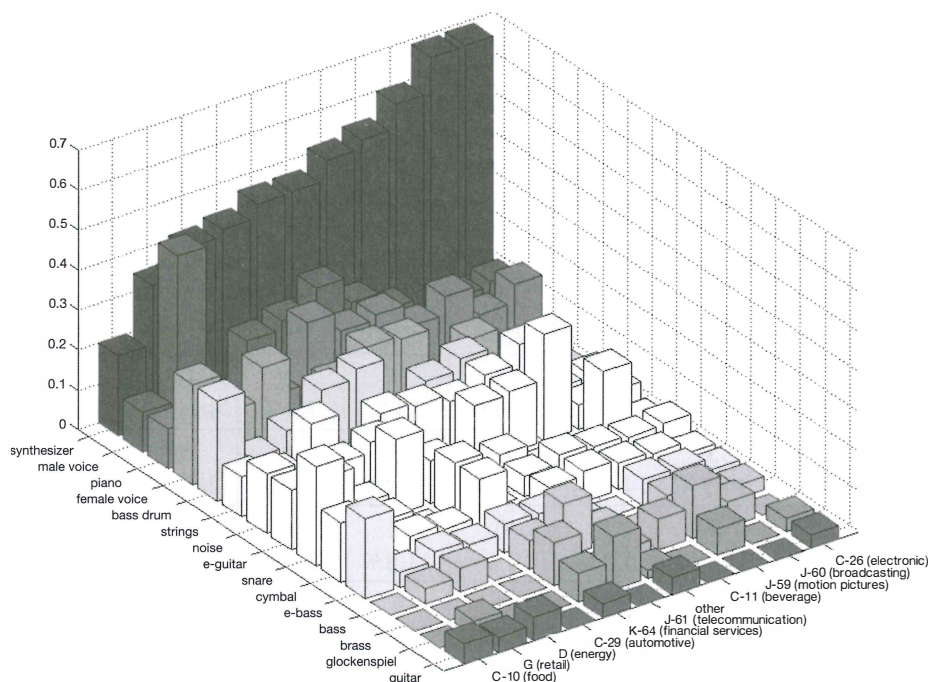


Abb. 1:
Relative Häufigkeiten vorkommender Instrumente in untersuchten Audiologos
verschiedener Branchen

Melodieverlauf, der auf einem geräuschhaft-perkussiven Klang beruht – vielleicht auch als akustisches Pendant zu einem gleichmäßigen, beständigen und sicheren Straßenverlauf (z. B. *Hyundai, Mazda, Audi, BMW*).

3.4 Stimme

Die Ausprägung der Stimme wird in den Audiologos (falls vorhanden) sehr unterschiedlich eingesetzt, sei es als Sprechstimme (spoken), singend (singing), flüsternd (whispered), pfeifend (whistled), summend (humming) oder schreiend (screaming), sei es platziert im Vordergrund oder in die Musik bzw. Geräusche eingebettet (inherent) oder von dieser bzw. diesen getrennt (additional), etwa als Claim. Anders als bei der Melodiekontur zeigt ein Chi-Quadrat-Test im Falle der Stimme, dass sich sowohl die Art ihrer Verwendung oder Nicht-Verwendung innerhalb der Audiologos in Abhängigkeit von der Branche unterscheidet ($\chi^2 [18, N=200] = 50.275; p = .0001; \phi = .501$) als auch in Abhängigkeit vom jeweiligen Kontinent ($\chi^2 [4, N=200] = 50.817; p = .003; \phi = .281$). Besonders im Bereich des Einzelhandels (G) ist die Stimme (meist als Sprechstimme) sehr präsent (88.5 %) (z. B. *Ebay, Edeka, Fressnapf, Obi*), vor allem bei europäischen Audiologos.

Tab. 3:

Anzahl und prozentualer Anteil der Audiologos unterteilt nach Kontur und Branche

| Industry | Contour | | | | | | Total |
|------------------------------------|----------------|-------------|--------------|-----------------|-----------------|---------------|--------------|
| | ascen- ding | concave | convex | descen- ding | hori- zontal | wave- like | |
| C-10 (food) | 3 15 % | 3 15 % | 3 15 % | 2 10 % | 1 5 % | 8 40 % | 20 100 % |
| C-11 (beverage) | 4 33.3 % | 1 8.3 % | 3 25 % | 0 0 % | 0 0 % | 4 33.3 % | 12 100 % |
| C-26 (electronic) | 7 25.9 % | 1 3.7 % | 3 11.1 % | 1 3.7 % | 2 7.4 % | 13 48.1 % | 27 100 % |
| C-29 (automotive) | 3 14.3 % | 1 4.8 % | 5 23.8 % | 3 14.3 % | 4 19 % | 5 23.8 % | 21 100 % |
| D (energy) | 9 52.9 % | 1 5.9 % | 0 0 % | 2 11.8 % | 1 5.9 % | 4 23.5 % | 17 100 % |
| G (retail) | 3 11.5 % | 5 19.2 % | 4 15.4 % | 7 26.9 % | 1 3.8 % | 6 23.1 % | 26 100 % |
| J-59 (motion pictures) | 5 33.3 % | 1 6.7 % | 1 6.7 % | 0 0 % | 1 6.7 % | 7 46.7 % | 15 100 % |
| J-60 (broad- casting) | 4 25 % | 0 0 % | 5 31.3 % | 1 6.3 % | 0 0 % | 6 37.5 % | 16 100 % |
| J-61 (tele- communica- tion) | 5 35.7 % | 1 7.1 % | 6 42.9 % | 2 14.3 % | 0 0 % | 0 0 % | 14 100 % |
| K-64 (financial services) | 5 17.9 % | 1 3.6 % | 4 14.3 % | 1 3.6 % | 1 3.6 % | 16 57.1 % | 28 100 % |
| Total | 48 24.5 % | 15 7.7 % | 34 17.3 % | 19 9.7 % | 11 5.6 % | 69 35.2 % | 196 100 % |

(82 %), ähnlich wie auch in der Telekommunikationsbranche (J-61; 53 %) (z. B. *AI, Bob, Congstar, Vodafone*). In Branchen, in denen es weniger auf die direkte Kommunikation ankommt, wie in der Getränke- (C11) und Elektronikindustrie (C-26), oder in denen das Audiologo eher eine Gesamtmarke oder -institution repräsentiert als ein einzelnes Produkt, wie in der Rundfunk- (J-60) und Filmbranche (J-59), findet auch die Stimme einen deutlich geringeren Einsatz (z. B. *Premiere, kabel bw*). Die Variationsvielfalt der Stimme wird am stärksten im Einzelhandel (G) und in der Lebensmittelindustrie (C-11) ausgenutzt: Hier dient die Stimme nicht nur der Informationsvermittlung und dem „Product Placement“, sondern sie wird auch in gesungener, gepfiffener (z. B. *Red Bull Mobile*),

geschriener (z.B. *Zalando*, *Saturn*, *Playstation 2*) und gesummter Form (z.B. *Exquisa*, *Weihenstephan*) eingesetzt, während in den Branchen der Energieanbieter (D), der Automobil- (C-29) und Elektronikindustrie (C-26) vereinzelt auch geflüstert wird (z.B. *Mazda*, *Seat*, *SANYO*) (vgl. Abb. 2 und Tab. 4).

3.5 Klangliche, strukturelle und psychoakustische Merkmale

Wenn sich die Audiologos allein schon vom Höreindruck anhand der musikalischen Merkmale wie Instrumentation, Melodiekontur und Einsatz der Stimme einzelnen Branchen zuordnen lassen, dann lässt die rechnerische Analyse der klanglichen Merkmale via MATLAB/MIRtoolbox noch weitaus genauere und objektivierbare Unterscheidungskriterien und Zuordnungsmöglichkeiten zu. So zeigte sich bei einer ANOVA über die analysierten (psycho-)akustischen Merkmale, dass sich die Branchen hinsichtlich der Parameter Spectral Centroid ($F [9, 190] = 3.396; p = .001$), Spectral Flux ($F [9, 190] = 1.955; p = .047$), abso-

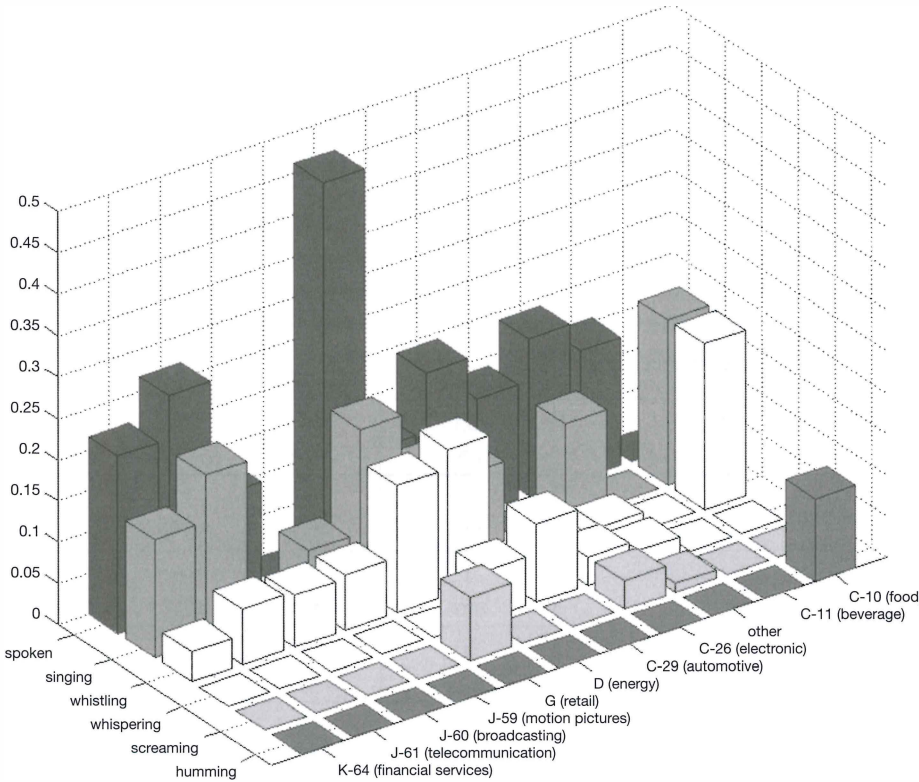


Abb. 2:
Relative Häufigkeiten des Einsatzes der Stimme in untersuchten Audiologos
verschiedener Branchen

Tab. 4:
Anzahl und prozentualer Anteil der Audiologos unterteilt nach Stimme und Branche
sowie Stimme und Kontinent

| | | Voice | | | Total |
|-----------|---------------------------|--------------|--------------|---------------|--------------|
| | | Additional | Inherent | No | |
| Branche | C-10 (food) | 2 10 % | 8 40 % | 10 50 % | 20 100 % |
| | C-11 (beverage) | 0 0 % | 2 15.4 % | 11 84.6 % | 13 100 % |
| | C-26 (electronic) | 3 10.3 % | 3 10.3 % | 23 79.3 % | 29 100 % |
| | C-29 (automotive) | 2 9.5 % | 6 28.6 % | 13 61.9 % | 21 100 % |
| | D (energy) | 4 23.5 % | 4 23.5 % | 9 52.9 % | 17 100 % |
| | G (retail) | 10 38.5 % | 13 50 % | 3 11.5 % | 26 100 % |
| | J-59 (motion pictures) | 0 0 % | 2 13.3 % | 13 86.7 % | 15 100 % |
| | J-60 (broadcasting) | 3 18.8 % | 0 0 % | 13 81.3 % | 16 100 % |
| | J-61 (telecommunication) | 3 20 % | 5 33.3 % | 7 46.7 % | 15 100 % |
| | K-64 (financial services) | 6 21.4 % | 5 17.9 % | 17 60.7 % | 28 100 % |
| | Total | 33 16.5 % | 48 24 % | 119 59.5 % | 200 100 % |
| Continent | America | 2 5 % | 6 15 % | 32 80 % | 40 100 % |
| | Asia | 1 5 % | 3 15 % | 16 80 % | 20 100 % |
| | Europe | 30 21.4 % | 39 27.9 % | 71 50.7 % | 140 100 % |
| | Total | 33 16.5 % | 48 24 % | 119 59.5 % | 200 100 % |

luter Dynamikumfang ($F [9, 190] = 2.533; p = .009$) und erste Einschwingzeit ($F [9, 190] = 2.814; p = .004$) signifikant unterscheiden (Signifikanzniveau $p < .05$) (vgl. Tab. 6). Die Parameter Inharmonizität, Rauigkeit, Niedrigenergieanteil, Unangenehmheit sowie relativer Dynamikumfang ergaben hingegen keine signifikanten Ergebnisse. Der Parameter Inharmonizität scheint jedoch tendenziell als Differenzierungskriterium geeignet zu sein ($F [9, 190] = 1.907; p = .054$).

In einer auf die ANOVA folgenden Post-hoc-Analyse (Tukeys HSD, vgl. Tab. 5) wurde deutlich, dass sich hinsichtlich der psychoakustischen Parameter besonders Audiologos aus der Filmindustrie (J-59) von den Audiologos der meisten anderen Gruppen unterscheiden. So zeigen sich Unterschiede in der ersten Einschwingzeit im Vergleich zur Lebensmittel- (C-10), Elektronik- (C-26) und Automobilindustrie (C-29), zu Energieanbietern (D), zum Einzelhandel (G) und Finanzdienstleistern (K-64) sowie Unterschiede hinsichtlich des Spectral Centroids im Vergleich zum Einzelhandel (G).

Während sich die Audiologos aus der Telekommunikationsbranche (J-61) von den Logos aller anderen Branchen in Bezug auf sämtliche untersuchten psychoakustischen Parameter nicht signifikant unterscheiden, finden sich weitere Unterschiede zwischen den Branchen der Getränkehersteller (C-11) und Energieanbieter (D) bezüglich des absoluten Dynamikumfangs, zwischen der Elektronikindustrie (C-26) und dem Einzelhandel (G) hinsichtlich des Spectral Centroids und des Spectral Flux sowie zwischen dem Einzelhandel (G) und dem Rundfunk (J-60) unter dem Aspekt des Spectral Centroids (vgl. Tab. 5 und 6).

Ähnliches gilt für die strukturellen Merkmale der Audiologos, die sich nach einer ANOVA der ermittelten Daten bezüglich der einzelnen Daten signifikant in der Länge ($F [9, 190] = 10.310; p = .0001$), der Anzahl der Ereignisse ($F [9, 90] = 1.958; p = .046$), der Anzahl der einzelnen Tönhöhensegmente ($F [9, 190] = 7.742; p = .0001$) und der Anzahl der harmonischen Segmente ($F [9, 190] = 8.497; p = .0001$) unterscheiden (vgl. Tab. 7 und 8).

Der Parameter „Anzahl der Ereignisse“ ergab im paarweisen Vergleich der anschließenden Post-hoc-Analyse (Tukeys HSD) keine signifikanten Ergebnisse.

Auffallend ist – wie schon bei den psychoakustischen Merkmalen – die Sonderstellung der Audiologos aus der Filmindustrie (J-59), die sich bezüglich der Merkmale Länge der Audiologos, Anzahl der Tönhöhensegmente sowie Anzahl der harmonischen Segmente von allen anderen Branchenzweigen deutlich unterscheiden (vgl. Tab. 8). Darüber hinaus korreliert die Länge der Audiologos signifikant mit der Dauer des ersten Einschwingvorgangs (Pearsons $r = .407; p = .0001$) sowie natürlich mit der Anzahl der Tönhöhensegmente (Pearsons $r = .860; p = .0001$) und der Anzahl der harmonischen Segmente (Pearsons $r = .927; p = .0001$).

Setzt man die ermittelten (psycho-)akustischen und strukturellen Merkmale über eine ANOVA mit anschließender Post-hoc-Analyse in eine Beziehung zum Kontinent, aus dem das Audiologo stammt bzw. wo es hauptsächlich verwendet wird, dann ergeben sich ebenfalls signifikante Unterschiede. So lassen sich anhand des Spectral Centroids ($F [2, 197] = 5.093; p = .007$), des Spectral Flux ($F [2, 197] = 4.670; p = .010$), der Inharmonizität ($F [2, 197] = 3.303; p = .039$) und

Tab. 5:

Deskriptive Statistik für die Parameter *Erste Einschwingzeit*, *Spectral Centroid*, *Spectral Flux* und *Absoluter Dynamikumfang* unterteilt nach Branchen

| Branche | N | Erste Einschwingzeit (in Sekunden) | | | Spectral Centroid | | | Spectral Flux | | | Absoluter Dynamik- umfang | | |
|------------------------------------|----|---------------------------------------|-------|------|-------------------|--------|-------|---------------|------|------|------------------------------|------|------|
| | | M | SD | SE | M | SD | SE | M | SD | SE | M | SD | SE |
| C-10 (food) | 20 | .262 | .150 | .034 | 2762.8 | 1300.6 | 290.8 | 108.2 | 52 | 11.6 | .403 | .122 | .027 |
| C-11 (beverage) | 13 | .429 | .533 | .148 | 3494.4 | 2012.5 | 558.2 | 97.5 | 52.3 | 14.5 | .282 | .071 | .020 |
| C-26 (electronic) | 29 | .344 | .357 | .066 | 2198.3 | 958.5 | 178 | 78.4 | 56.8 | 10.5 | .342 | .120 | .022 |
| C-29 (automotive) | 21 | .245 | .150 | .033 | 2852.1 | 1400.2 | 305.5 | 106.1 | 46.9 | 10.2 | .388 | .123 | .027 |
| D (energy) | 17 | .303 | .450 | .109 | 2833.7 | 1550.1 | 376 | 103.9 | 77.9 | 18.9 | .436 | .115 | .028 |
| G (retail) | 26 | .210 | .104 | .020 | 3664.7 | 1319.2 | 258.7 | 137.1 | 74.6 | 14.6 | .347 | .103 | .020 |
| J-59 (motion pictures) | 15 | .907 | 1.352 | .349 | 2191.2 | 1168.4 | 301.7 | 80.1 | 38.9 | 10 | .346 | .101 | .026 |
| J-60 (broad- casting) | 16 | .512 | .529 | .132 | 1902.9 | 1079.5 | 269.9 | 87.6 | 42.3 | 10.6 | .404 | .133 | .033 |
| J-61 (tele- communica- tion) | 15 | .345 | .346 | .089 | 2464.1 | 1655.3 | 427.4 | 123.5 | 79.9 | 20.6 | .419 | .157 | .040 |
| K-64 (financial services) | 28 | .272 | .391 | .074 | 2648.8 | 1253.9 | 237 | 91.1 | 78.8 | 14.9 | .351 | .110 | .021 |

Anmerkung: N: Anzahl der Audiologos, M: Mittelwert, SD: Standardabweichung, SE: Standardfehler

Tab. 6:

Paarweiser Mittelwertvergleich (Tukeys HSD) der *ersten Einschwingzeit*, des *Spectral Centriods*, des *Spectral Flux* und des *absoluten Dynamikumfangs* zwischen den verschiedenen Branchen

| | Branche A | Branche B | <i>p</i> | <i>SE</i> | <i>C.I.</i>_{LB} | <i>C.I.</i>_{UB} | <i>Cohens d</i> |
|---|---------------------------|---------------------------|-----------------|------------------|---------------------------------|---------------------------------|------------------------|
| Erste Einschwingzeit (in Sekunden) | J-59 (motion pictures) | C-10 (food) | .007 | .17 | .10 | 1.19 | .67 |
| | | C-26 (electronic) | .016 | .16 | .06 | 1.07 | .57 |
| | | C-29 (auto-motive) | .004 | .17 | .12 | 1.20 | .69 |
| | | D (energy) | .025 | .03 | .04 | 1.17 | .60 |
| | | G (retail) | .001 | .01 | .18 | 1.21 | .73 |
| | | K-64 (financial services) | .004 | .01 | .13 | 1.14 | .64 |
| Spectral Centriods | G (retail) | J-59 (motion pictures) | .030 | 436.9 | 74.5 | 2872.4 | 1.18 |
| | | C-26 (electronic) | .003 | 364.0 | 301.1 | 2631.7 | 1.27 |
| | | J-60 (broadcasting) | .002 | 428.2 | 390.9 | 3132.7 | 1.46 |
| Spectral Flux | G (retail) | C-26 (electronic) | .024 | 17.08 | 4.0 | 113.4 | .89 |
| Absoluter Dynamikumfang | D (energy) | C-11 (beverage) | .016 | .04 | .02 | .29 | 1.61 |

Anmerkung: Es werden lediglich die signifikanten Ergebnisse dargestellt (Signifikanzniveau $p < .05$).
p: *p*-Wert, *SE:* Standardfehler, *C.I.*_{LB}: Untergrenze Konfidenzintervall (95 %), *C.I.*_{UB}: Obergrenze Konfidenzintervall (95 %)

Tab. 7:

Deskriptive Statistik für die Parameter *Länge der Audiologos*, *Anzahl der Ereignisse*, *Anzahl der Tonhöhensegmente* und *Anzahl der harmonischen Segmente* unterteilt nach Branchen

| Branche | N | Länge der Audiologos (in Sekunden) | | | Anzahl der Ereignisse | | | Anzahl der Tonhöhensegmente | | | Anzahl der harmonischen Segmente | | |
|---------------------------------|----|---------------------------------------|------|------|-----------------------|------|------|--------------------------------|-------|------|-------------------------------------|-------|------|
| | | M | SD | SE | M | SD | SE | M | SD | SE | M | SD | SE |
| C-10 (food) | 20 | 5.69 | 4.13 | .92 | 7.35 | 5.71 | 1.28 | 13.85 | 10.59 | 2.37 | 10.45 | 8.23 | 1.84 |
| C-11 (beverage) | 13 | 6.75 | 4.59 | 1.27 | 7.85 | 5.60 | 1.55 | 14.62 | 9.05 | 2.51 | 11.77 | 7.92 | 2.20 |
| C-26 (electronic) | 29 | 4.87 | 3.46 | .64 | 5.21 | 3.72 | .69 | 11.03 | 7.08 | 1.31 | 8.72 | 7.63 | 1.42 |
| C-29 (automotive) | 21 | 3.04 | 1.25 | .27 | 4.33 | 1.71 | .37 | 7.14 | 2.74 | .60 | 5.38 | 1.77 | .39 |
| D (energy) | 17 | 4.39 | 1.95 | .47 | 5.00 | 2.55 | .62 | 8.94 | 3.15 | .76 | 7.65 | 3.26 | .79 |
| G (retail) | 26 | 2.63 | 1.11 | .22 | 5.96 | 2.75 | .54 | 7.15 | 2.59 | .51 | 5.19 | 2.14 | .42 |
| J-59 (motion pictures) | 15 | 11.13 | 5.48 | 1.41 | 6.60 | 5.44 | 1.40 | 24.07 | 14.78 | 3.82 | 20.00 | 10.45 | 2.70 |
| J-60 (broad- casting) | 16 | 5.82 | 3.53 | .88 | 4.81 | 2.43 | .61 | 11.13 | 7.87 | 1.97 | 10.19 | 7.43 | 1.86 |
| J-61 (telecom- munication) | 15 | 3.14 | 1.79 | .46 | 4.53 | 1.77 | .46 | 8.07 | 4.64 | 1.20 | 5.53 | 2.59 | .67 |
| K-64 (financial services) | 28 | 4.95 | 2.53 | .48 | 6.82 | 3.81 | .72 | 11.11 | 6.17 | 1.17 | 9.18 | 4.72 | .89 |

Anmerkung: N: Anzahl der Audiologos, M: Mittelwert, SD: Standardabweichung, SE: Standardfehler

Tab. 8:
Paarweiser Mittelwertvergleich (Tukeys HSD) der *Länge der Audiologos*, der *Anzahl der Tonhöhensegmente* und der *Anzahl der harmonischen Segmente* zwischen den verschiedenen Branchen

| | Branche A | Branche B | <i>p</i> | <i>SE</i> | <i>C.I.</i> _{LB} | <i>C.I.</i> _{UB} | Cohens <i>d</i> |
|---------------------------------------|---------------------------|----------------------|----------|-----------|---------------------------|---------------------------|-----------------|
| Länge der Audiologos (in Sekunden) | J-59 (motion pictures) | C-10 (food) | .0001 | 1.069 | 2.019 | 8.865 | 1.12 |
| | | C-11 (beverage) | .0110 | 1.186 | .580 | 8.175 | .87 |
| | | C-26 (electronic) | .0001 | .995 | 3.076 | 9.450 | 1.37 |
| | | C-29 (automotive) | .0001 | 1.058 | 4.703 | 11.478 | 2.04 |
| | | D (energy) | .0001 | 1.109 | 3.185 | 10.285 | 1.64 |
| | | G (retail) | .0001 | 1.015 | 5.251 | 11.749 | 2.15 |
| Anzahl der Tonhöhensegmente | J-59 (motion pictures) | C-10 (food) | .0030 | 2.516 | 2.16 | 18.27 | .79 |
| | | C-11 (beverage) | .0290 | 2.792 | .51 | 18.39 | .77 |
| | | C-26 (electronic) | .0001 | 2.343 | 5.53 | 20.53 | 1.12 |
| | | C-29 (automotive) | .0001 | 2.490 | 8.95 | 24.90 | 1.59 |
| | | D (energy) | .0001 | 2.610 | 6.77 | 23.48 | 1.42 |
| | | G (retail) | .0001 | 2.389 | 9.27 | 24.56 | 1.59 |
| | | J-60 (broad-casting) | .0010 | 2.648 | 4.46 | 21.42 | 1.09 |
| | | J-61 (telecom.) | .0001 | 2.690 | 7.39 | 24.61 | 1.46 |
| | | K-64 (fin. serv.) | .0001 | 2.357 | 5.41 | 20.51 | 1.14 |

Tab. 8:
Fortsetzung

| | Branche A | Branche B | <i>p</i> | <i>SE</i> | <i>C.I.</i> _{LB} | <i>C.I.</i> _{UB} | Cohens <i>d</i> |
|----------------------------------|---------------------------|----------------------|----------|-----------|---------------------------|---------------------------|-----------------|
| Anzahl der harmonischen Segmente | J-59 (motion pictures) | C-10 (food) | .0001 | 2.084 | 2.88 | 16.22 | 1.02 |
| | | C-11 (beverage) | .0160 | 2.312 | .83 | 15.63 | .89 |
| | | C-26 (electronic) | .0001 | 1.94 | 5.06 | 17.49 | 1.23 |
| | | C-29 (automotive) | .0001 | 2.062 | 8.02 | 21.22 | 1.95 |
| | | D (energy) | .0001 | 2.161 | 5.43 | 19.27 | 1.60 |
| | | G (retail) | .0001 | 1.978 | 8.48 | 21.14 | 1.96 |
| | | J-60 (broad-casting) | .0010 | 2.192 | 2.79 | 16.83 | 1.08 |
| | | J-61 (telecom.) | .0001 | 2.227 | 7.33 | 21.6 | 1.90 |
| | | K-64 (fin. serv.) | .0001 | 1.952 | 4.57 | 17.07 | 1.33 |

Anmerkung: Es werden lediglich die signifikanten Ergebnisse dargestellt (Signifikanzniveau $p < .05$).
p: *p*-Wert, *SE*: Standardfehler, *C.I.*_{LB}: Untergrenze Konfidenzintervall (95 %), *C.I.*_{UB}: Obergrenze Konfidenzintervall (95 %)

der ersten Einschwingzeit ($F [2, 197] = 5.369$; $p = .005$) die Audiologos klanglich den einzelnen Kontinenten zuordnen (vgl. Tab. 9 und 10; da Audiologos mit der Branchenzuordnung „Entertainment“ häufig auch aus Amerika stammen, muss die geografische Unterscheidbarkeit hier vielleicht etwas relativiert werden).

Ebenso lassen sich die Länge der Audiologos ($F [2, 197] = 16.335$; $p = .0001$), die Anzahl der Ereignisse ($F [2, 197] = 3.860$; $p = .023$), die Anzahl der einzelnen Tonhöhensegmente ($F [2, 197] = 16.138$; $p = .0001$) sowie die Anzahl der harmonischen Segmente ($F [2, 197] = 13.245$; $p = .0001$) zur signifikanten geografischen Unterscheidung aufgrund klanglicher Merkmale heranziehen (vgl. Tab. 11 und 12).

Trägt man all diese als signifikant ermittelten geografischen, musikalischen und psychoakustischen Unterschiede zwischen den Branchen in einer Matrix gegeneinander auf, so ergibt sich schnell ein Gesamtüberblick über die Parameter, anhand derer sich die Audiologos der einzelnen Branchen allein aufgrund ihrer klanglichen Merkmale unterscheiden lassen (vgl. Abb. 3): Es zeigt sich deutlich, dass sich die Audiologos aus der Filmbranche (J-59) besonders durch Parameter wie erste Einschwingzeit (in der Matrix: AT), Länge (in der Matrix: L) und die Anzahl der tonalen (in der Matrix: PS) und harmonischen Segmente (in der Matrix: HS) von den anderen Audiologos unterscheiden, während die Audiologos des Einzelhandels sich global betrachtet meist durch den Einsatz

| | C-10 (food) | C-11 (beverage) | C-26 (electronic) | C-29 (automotive) | D (energy) | G (retail) | J-59 (motion pictures) | J-60 (broadcasting) | J-61 (telecommu- nication) | K-64 (financial services) |
|----------------------------------|----------------|--------------------|----------------------|----------------------|---------------|--------------------------|---------------------------------|-------------------------|----------------------------------|---------------------------------|
| C-10 (food) | | | M Cnt - Co | M | | M - V L | AT - L - PS - HS Cnt - Co | | | |
| C-11 (beverage) | | | Cnt - Co | L | Dyn | V L | L - PS - HS Cnt - Co | | | |
| C-26 (electronic) | | | | Cnt - Co | Co | V SC - SF Cnt - Co | AT - L - PS - HS | | | Cnt - Co |
| C-29 (automotive) | | | | | | V | AT - L - PS - HS Cnt - Co | | | |
| D (energy) | | | | | | | AT - L - PS - HS Cnt - Co | | | |
| G (retail) | | | | | | | V SC - AT - HS Cnt - Co | V SC - L | | V |
| J-59 (motion pictures) | | | | | | | | L - PS - HS Cnt - Co | L - PS - HS Cnt - Co | AT - L - PS - HS Cnt - Co |
| J-60 (broadcasting) | | | | | | | | | | |
| J-61 (telecommuni- cation) | | | | | | | | | | Cont |
| K-64 (financial services) | | | | | | | | | | |

Abb. 3:

Signifikante branchentypische Unterschiede in den klanglichen Eigenschaften von Audiologos im Überblick: *Musikalische Eigenschaften:* M=Manner, V=Voice, Cont=Contour. *Psychoakustische Eigenschaften:* AT=First Attack Time, L=Length, Dyn=abs. Dynamic Range (all), SC=Spectral Centroid, SF=Spectral Flux (mean and median), PS=Num Pitch Segments, HS=Num Harmonic Segments. *Geographische Eigenschaften:* Cnt=Continent, Co=Country

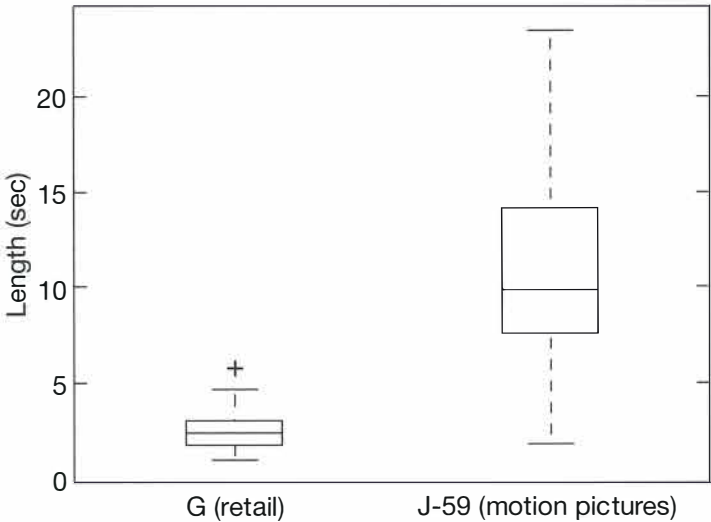


Abb. 4:

Paarweiser Längenvergleich (in Sekunden) von Audiologos aus dem Einzelhandel (G) vs. Audiologos aus der Filmbranche (J-59)

der Stimme (in der Matrix: V) von den Audiologos der anderen Branchen abheben. Auch die Audiologos einzelner Branchen lassen sich anhand psychoakustischer Merkmale wie Spectral Centroid (in der Matrix: SC) und Spectral Flux (in der Matrix: SF) deutlich voneinander unterscheiden (z. B. Einzelhandel [G] vs. Elektronikbranche [C-26] oder Einzelhandel [G] vs. Rundfunk [J-60]), ebenso wie durch die Art und Weise (in der Matrix: M), wie die Sing- und Sprechstimme eingesetzt wird (z. B. Lebensmittel- [C-11] vs. Elektronikbranche [C-26] oder Automobilindustrie [C-29]) oder wie die Kontur (in der Matrix: Cont) der Melodieführung gestaltet ist (z. B. Telekommunikation [J-61] vs. Finanzdienstleistungsbranche [K-64]) (vgl. Abb. 4).

Tab. 9:
Deskriptive Statistik für die Parameter *Erste Einschwingzeit*, *Spectral Centroid*, *Spectral Flux*, *Inharmonizität* unterteilt nach Kontinenten

| | Continent | N | M | SD | SE |
|--|-----------|-----|--------|--------|-------|
| Erste Einschwingzeit (in Sekunden) | America | 40 | .589 | .914 | .145 |
| | Asia | 20 | .324 | .400 | .089 |
| | Europe | 140 | .293 | .327 | .028 |
| Spectral Centroid | America | 40 | 2151.8 | 1156.5 | 182.9 |
| | Asia | 20 | 2430.0 | 1318.0 | 294.7 |
| | Europe | 140 | 2911.1 | 1458.5 | 123.3 |
| Spectral Flux | America | 40 | 76.5 | 45.3 | 7.2 |
| | Asia | 20 | 90.4 | 50.7 | 11.3 |
| | Europe | 140 | 110.0 | 69.1 | 5.8 |
| Inharmonizität | America | 40 | .446 | .057 | .009 |
| | Asia | 20 | .398 | .138 | .031 |
| | Europe | 140 | .436 | .059 | .005 |

Anmerkung: N: Anzahl der Audiologos, M: Mittelwert, SD: Standardabweichung, SE: Standardfehler

Tab. 10:
Paarweiser Mittelwertvergleich (Tukey’s HSD) der *Ersten Einschwingzeit*,
des *Spectral Centriods*, *Spectral Flux* und der *Inharmonizität* zwischen den
verschiedenen Kontinenten

| | Continent A | Continent B | <i>p</i> | <i>SE</i> | <i>C.I.</i> _{LB} | <i>C.I.</i> _{UB} | Cohens <i>d</i> |
|---------------------------------|----------------|----------------|----------|-----------|---------------------------|---------------------------|--------------------|
| Erste Ein- schwing- zeit | Europe | America | .004 | .094 | .082 | .510 | .43 |
| Spectral Centriods | Europe | America | .007 | 249.3 | 170.6 | 1348.0 | .58 |
| Spectral Flux | Europe | America | .010 | 11.4 | 6.6 | 60.3 | .42 |
| Absoluter Dynamik- umfang | Asia | America | .035 | .019 | .003 | .094 | .46 |

Anmerkung: Es werden lediglich die signifikanten Ergebnisse dargestellt (Signifikanzniveau $p < .05$).
p: *p*-Wert, *SE*: Standardfehler, *C.I.*_{LB}: Untergrenze Konfidenzintervall (95 %), *C.I.*_{UB}:
Obergrenze Konfidenzintervall (95 %)

Tab. 11:
Deskriptive Statistik für die Parameter *Länge der Audiologos*, *Anzahl der Ereignisse*,
Anzahl der Tonhöhensegmente und *Anzahl der harmonischen Segmente* unterteilt
nach Kontinenten

| | Continent | <i>N</i> | <i>M</i> | <i>SD</i> | <i>SE</i> |
|--|-----------|----------|----------|-----------|-----------|
| Länge der Audiologos (in Sekunden) | America | 40 | 7.764 | 5.250 | .830 |
| | Asia | 20 | 3.787 | 2.716 | .607 |
| | Europe | 140 | 4.349 | 2.890 | .244 |
| Anzahl der Ereignisse | America | 40 | 5.15 | 4.073 | .644 |
| | Asia | 20 | 4.05 | 2.564 | .573 |
| | Europe | 140 | 6.28 | 3.827 | .323 |

Tab. 11:
Fortsetzung

| | Continent | <i>N</i> | <i>M</i> | <i>SD</i> | <i>SE</i> |
|---------------------------------------|-----------|----------|----------|-----------|-----------|
| Anzahl der Tonhöhen-segmente | America | 40 | 17.48 | 12.594 | 1.991 |
| | Asia | 20 | 8.45 | 5.286 | 1.182 |
| | Europe | 140 | 9.84 | 6.229 | .526 |
| Anzahl der harmo- nischen Segmente | America | 40 | 13.80 | 10.481 | 1.657 |
| | Asia | 20 | 7.00 | 5.629 | 1.259 |
| | Europe | 140 | 7.89 | 5.266 | .445 |

Anmerkung: *N*: Anzahl der Audiologos, *M*: Mittelwert, *SD*: Standardabweichung, *SE*: Standardfehler

Tab. 12:

Paarweiser Mittelwertvergleich (Tukeys HSD) der *Länge der Audiologos*, der *Anzahl der Ereignisse*, der *Anzahl der Tonhöhen-segmente* und der *Anzahl der harmonischen Segmente* zwischen den verschiedenen Kontinenten

| | Continent A | Continent B | <i>p</i> | <i>SE</i> | <i>C.I.</i> _{LB} | <i>C.I.</i> _{UB} | Cohen's <i>d</i> |
|---|-------------|-------------|----------|-----------|---------------------------|---------------------------|---------------------|
| Länge der Audio- logos (in Sekunden) | America | Asia | .0001 | .951 | 1.730 | 6.223 | .95 |
| | | Europe | .0001 | .623 | 1.944 | 4.885 | .81 |
| Anzahl der Ereignisse | Europe | Asia | .038 | .903 | .10 | 4.36 | .68 |
| Anzahl der Tonhöhen- segmente | America | Asia | .0001 | 2.147 | 3.95 | 14.10 | .93 |
| | | Europe | .0001 | 1.406 | 4.31 | 10.95 | .77 |
| Anzahl der harmoni- schen Segmente | America | Asia | .0010 | 1.824 | 2.49 | 11.11 | .81 |
| | | Europe | .0001 | 1.194 | 3.09 | 8.73 | .71 |

Anmerkung: Es werden lediglich die signifikanten Ergebnisse dargestellt (Signifikanzniveau $p < .05$).
p: *p*-Wert, *SE*: Standardfehler, *C.I.*_{LB}: Untergrenze Konfidenzintervall (95 %), *C.I.*_{UB}: Ober-
grenze Konfidenzintervall (95 %)

Mit anderen Worten: Aus den in der Audiologo-Branchen-Matrix zusammengestellten Ergebnissen wird es (ebenso wie aus den Tabellen 5 bis 8) klar ersichtlich, dass sich Audiologos allein schon aufgrund ihrer klanglichen Merkmale in vielen Fällen signifikant branchentypisch einordnen und unterscheiden lassen.

4 Diskussion

Die eingangs gestellte Frage, ob es bestimmte Klangeigenschaften gibt, die besonders typisch für Audiologos sind, lässt sich mit „Ja“ beantworten. Es muss allerdings betont werden, dass die diskutierten Klangeigenschaften zu einem großen Teil auf einer psychoakustischen Analyse basieren, die nicht unbedingt analog mit der vom Menschen wahrgenommenen Klangfarbe als einem Resultat komplexer Hörgestalten übereinstimmen muss.

Unabhängig von ihrer Branchenzugehörigkeit lassen sich als typische klangliche Merkmale von Audiologos synthetisch erzeugte Klänge mit einem Spectral Centroid im Bereich von zwei bis vier kHz finden, die über eine Dauer von ca. fünf Sekunden eine kurze wellenförmige oder aufsteigende Melodie bilden (vgl. auch Anzenbacher, 2012). Mit den ausgeprägten Frequenzanteilen im Bereich der Außenohreigenresonanz (2 bis 4 kHz) kann dabei ein aufmerksamkeitssteigernder und die Sprachverständlichkeit verbessernder Effekt einhergehen, der jedoch bei stärkeren Intensitäten auch leicht ins Unangenehme kippen kann (Reuter et al., 2014, S. 97). Das birgt dann die Gefahr, dass das Audiologo bei größeren Lautstärken dann eher als störend empfunden wird.

Schaut man sich die Audiologos eher unter branchenspezifischen Gesichtspunkten an, so wird deutlich, dass sich die klangliche und strukturelle Ausprägung der akustischen Marken- und Produktsymbole je nach Wirtschaftszweig deutlich unterscheiden kann: Besonders im Einzelhandel (G) wird bei den (im Durchschnitt 2.6 Sekunden kurzen) Audiologos eine Menge aufmerksamkeits-erregender Techniken eingesetzt, wie z. B. der häufige Einsatz der Stimme in ihren verschiedensten Ausprägungen, kurze Einschwingzeiten, signalähnliche Melodiekonturen sowie ein vergleichsweise hoher spektraler Schwerpunkt (Spectral Centroid) im Bereich von knapp 4000 Hz (z. B. bei *Zalando*, *real* oder *Reno*). Auch wenn die Häufigkeit sowie die Anwendungsvielfalt der Stimme im Einzelhandel (G) relativ hoch erscheinen, ist die menschliche Stimme als Hauptinformationsträger im Audiologo kein deutliches Merkmal zur Branchendifferenzierung. Auch das Merkmal, dass ein Audiologo häufig zusätzlich durch einen gesprochenen bzw. gesungenen Markenclaim erweitert wird (z. B. *Edeka*, *Fressnapf*, *Obi*, *Reno*, *Toom*) findet sich auch in den Fällen der anderen Branchen wieder. Die Stimme, so scheint es, wird als emotional wirksames akustisches Medium eingesetzt, um die Bereitschaft zur Informationsaufnahme auf Seite der Rezipienten zu steigern (Lehmann, 2016). Hinsichtlich der Informationsverarbeitung kann ein Sprachanteil dann nützlich sein, wenn die Annahme einer hohen Motivation und Fähigkeit zur Informationsaufnahme besteht, wie es anhand der zentralen Verarbeitung im Elaboration-Likelihood-Model von Petty und Cacioppo (1986) erklärt wird. Demnach können gesprochene Botschaften

intensiver verarbeitet werden und eine Einstellungsänderung gegenüber dem beworbenen Produkt oder der Marke beim Rezipienten erzielt werden. Bei geringem Involvement erhöht sich die Relevanz von anderen Hinweisreizen. Positive Effekte durch musikalische Elemente konnten bei der Verarbeitung einer Mitteilung und dem Aufbau einer emotionalen Beziehung unter geringem Involvement beobachtet werden (Alpert et al., 2005; MacInnis & Park, 1991). So können musikalische Elemente ein geeignetes Mittel für den Bedeutungstransfer sein – solange sie ein Mindestmaß an Kongruenz zum beworbenen Produkt aufweisen (Kellaris et al., 1993).

Im Rahmen von Imagetransferstrategien (Aichner, 2014) wird die Stimme eingesetzt einerseits zum Ausdruck der individuellen Persönlichkeit, andererseits auch als Hinweis auf die Herkunft. So sollen die Audiologos aus der Automobilbranche, wie VW („Das Auto“), *Renault* („Créateur d’automobiles) oder *Seat* („Auto emoción“) den Country-of-Origin-Effekt nutzen und die Assoziationen zum Landesimage transportieren, und zwar auch außerhalb des jeweiligen Herkunftslands (u. a. Badri, Davis & Davis, 1995). Eine Metaanalyse von Peterson und Jolibert (1995) konnte größere Country-of-Origin-Effekte bei verbalen Produktbeschreibungen feststellen, als dies bei alleiniger Präsenz eines Produkts der Fall ist. Während also Textinhalte auf den Mehrwert eines Markenprodukts hinweisen können, sichern das Charisma in der Stimme sowie stimmliche Eigenheiten den Wiedererkennungseffekt beim Hörer und provozieren bei ihnen emotionale Reaktionen, die dann über Gefallen und Nichtgefallen, die Überzeugungskraft schlüssiger Argumentation und schließlich zur Entscheidungsbildung führen können (North et al., 2004). Wie im Werbekontext beobachtet werden konnte, gelingt es durch musikalische Elemente die Aufmerksamkeit auf die Botschaft zu lenken, den Rezipient zu motivieren, die Botschaft zu verarbeiten und das Kaufpotenzial zu erleichtern (Brooker & Wheatley, 1994; Kellaris & Kent, 1991; Kellaris et al., 1993; Yalch, 1991).

Beinhaltet ein Audiologo sowohl gesprochene als auch instrumentale Bestandteile, so könnte das die Ansprache zweier unterschiedlicher Verarbeitungsmodi (Elaboration-Likelihood-Model) begünstigen. Ist der Rezipient zur sorgfältigen Verarbeitung von Argumenten motiviert und fähig, sind textbasierte Markenargumente demzufolge nicht hinderlich. Ist die intensive Verarbeitung von Argumenten gering, dann kann der Bedeutungstransfer aber auch mittels klanglicher Hinweisreize erfolgen.

Besonders im Einzelhandel scheint es obligat, durch eine möglichst direkte zwischenmenschliche Kommunikation eine enge emotionale Beziehung und damit ein Vertrauensverhältnis aufzubauen und den potenziellen Kunden durch ein gesteigertes emotionales Involvement bei der Kaufentscheidung positiv zu beeinflussen und sich klar von der Massenkommunikation abzuheben. Diese Art der direkten Kommunikation wird auch als Face-to-Face-Marketing bezeichnet (Poth, 2001; Poth & Poth, 1999, S. 105). Dies könnte auch ein möglicher Hinweis dafür sein, warum serviceorientierte Branchen in der akustischen Markenkommunikation besonders häufig die menschliche Stimme einsetzen, und so mit dieser auch klanglich auf ein direktes und nachhaltiges Vertrauensverhältnis zum Kunden hinarbeiten.

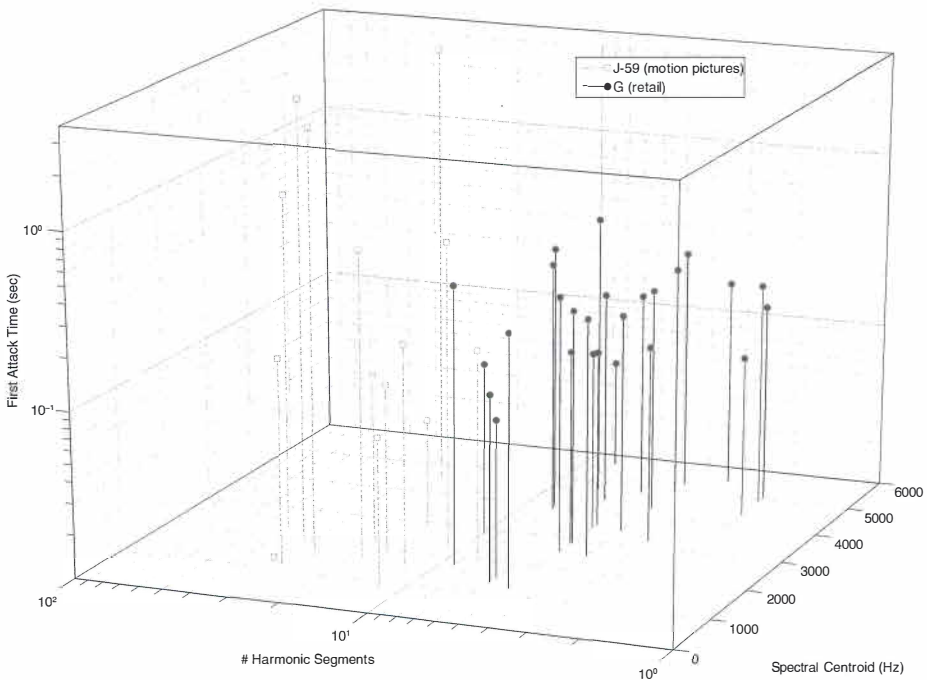
Ähnliches gilt auch für den gehäuftten Einsatz des Klaviers, welches als klanglicher Botschafter mit vertrautem, ruhigem und intelligentem Charakter (Bruner, 1990, S. 97) vor allem in Audiologos von Finanzdienstleistern (K-64) erklingt (wie z. B. bei *Allianz*, *DKSH*, *EasyCredit*, *Erste Bank*, *HUK Coburg*, *UBS*).

Den Audiologos des Einzelhandels (G) und der Automobilindustrie (C-29) klanglich diametral gegenübergestellt, finden sich die akustischen Markenklänge der Filmindustrie (J-59) und des Rundfunks (J-60), die häufig durch lange Einschwingzeiten, starke geräuschhafte Anteile, einem eher geringeren Spectral Centroid von etwas weniger als 2000 Hz und einem besonders im tieffrequenten Bereich modulierenden Sound gekennzeichnet sind (z. B. bei *DTS*, *Dreamcast* oder *Legendary Pictures*). Da die klangliche Gestaltung dieser Audiologos häufig synchron zum Verlauf eines bildhaften Logos oder Video-Intros verläuft, sind die Audiologos aus der Rundfunk- (J-60) und Filmbranche (J-59) auch häufig klanglich aufwendigere, vielfältigere und vor allem auch längerandauernde Klangproduktionen, was sich auch im direkten Längenvergleich zu den Audiologos aus dem Einzelhandel (G) zeigt.

Die gleich auf mehreren Dimensionen vorhandene klangliche Gegensätzlichkeit der Audiologos aus der Filmbranche (J-59) im Vergleich zu denen aus dem Einzelhandel (G) wird auch sofort erkennbar und offensichtlich, wenn man diese in einem Audiologo-Timbre-Space⁶ mit den Achsen „Anzahl der harmonischen Abschnitte“ (X-Achse: numHarmonic Segments), „erster Einschwingvorgang“ (Y-Achse: First Attack Time) und „spektraler Schwerpunkt“ (Z-Achse: Spectral Centroid) gegenüberstellt (es muss hierbei angemerkt werden, dass die räumliche Anordnung der branchentypischen Klangmerkmale keine Folgerungen auf eine eventuell subjektiv erlebte Ähnlichkeit der Audiologos innerhalb einer Branche bzw. zwischen den Branchen zulässt) (vgl. Abb. 5).

Nicht ganz so deutlich, aber ebenfalls sichtbar vorhanden, gestalten sich die klanglichen Unterschiede zwischen den Audiologos aus dem Einzelhandel (G) und denen aus der Elektronikbranche (C-26): In einem Audiologo-Timbre-Space mit den Dimensionen „Spektraler Schwerpunkt“ (X-Achse: Spectral Centroid), „erste Einschwingzeit“ (Y-Achse: First Attack Time) und „spektrale Fluktuationen“ (Z-Achse: Spectral Flux) lassen sich ebenfalls zwei räumliche Schwerpunkte deutlich erkennen, in denen auf der einen Seite die Audiologos der Elektronikbranche (C-26) mit geringerem spektralen Schwerpunkt, längeren Einschwingzeiten und weniger spektralen Fluktuationen versammelt sind und

6 Der Timbre-Space ist eine von John Grey 1975 eingeführte Klangfarbendarstellungsform, innerhalb derer Instrumentenklänge als Punkte entlang dreier – bestimmten Klangeigenschaften entsprechenden – Raumachsen angeordnet werden, um der Multidimensionalität von Klangfarbe gerecht zu werden. Während der Timbre-Space in seiner originalen Form für die Darstellung von Instrumentalklangfarben aufgrund mangelnder Verallgemeinerbarkeit und methodischer Mängel eher ungeeignet ist (vgl. Siddiq et al., 2015), hat er sich für die plastische Darstellung von klanglichen Unterscheidungskriterien als ein durchaus nützliches und sinnvolles Werkzeug erwiesen (z. B. Anzenbacher et al., 2015).

**Abb. 5:**

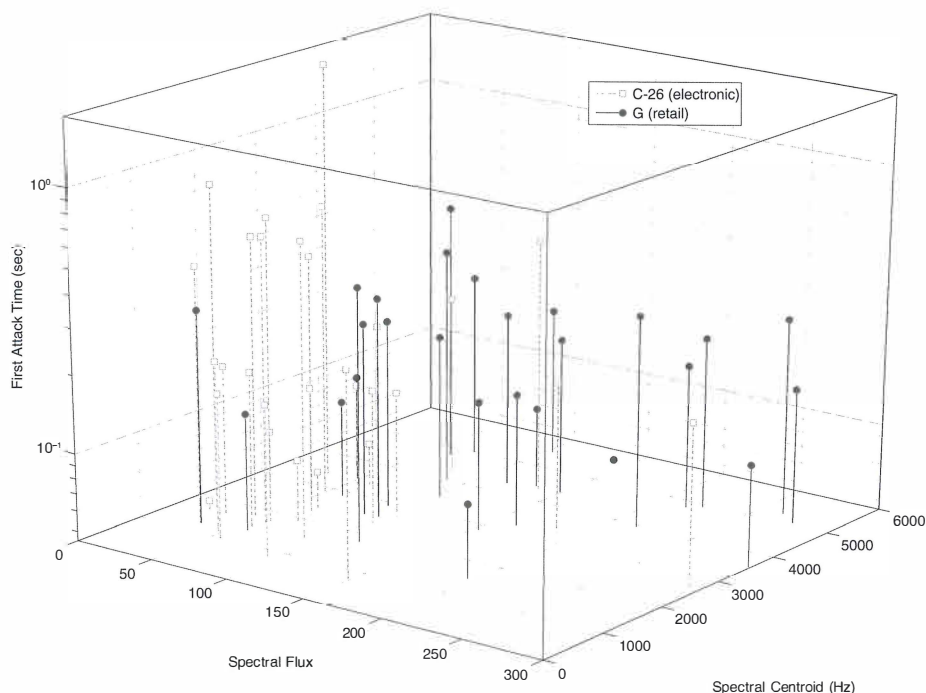
Audiologo-Timbre-Space für Audiologos aus dem Einzelhandel (G) vs. denen aus der Filmbranche (J-59) mit den Dimensionen „Number of Harmonic Segments“ (absolut), „First Attack Time“ (Sekunden) und „Spectral Centroid“ (Hz)

auf der anderen Seite die Audiologos des Einzelhandels (G) mit einem höheren spektralen Schwerpunkt, kürzeren Einschwingzeiten und (aufgrund der häufigeren Verwendung von Sing- und Sprechstimme) stärkeren spektralen Fluktuationen (vgl. Abb. 6).

Häufig lassen sich Audiologos aus verschiedenen Branchen auch schon allein anhand eines einzigen klanglichen Merkmals voneinander unterscheiden, wie z. B. beim Vergleich der verwendeten Kanaldynamik von Audiologos aus der Getränkeindustrie (C-11) mit denen aus der Energiebranche (D): Es wird schnell deutlich, dass in den Audiologos der Getränkeindustrie (C-11) eine deutlich stärkere Pegelkompression verwendet wird als bei denen in der Energiebranche (D), was allerdings auch an unterschiedlichen Produktionsstandards im Gestaltungs- und Implementierungsprozess der Audiologos liegen kann (vgl. Abb. 7).

Ein weiteres Unterscheidungskriterium zwischen den Audiologos dieser beiden Branchen ist der Einsatz der weiblichen Stimme, die in der Audiowerbung der Getränkeindustrie so gut wie nie verwendet wird, in den Audiologos der Energiebranche jedoch vergleichsweise häufig (vgl. Kap. 3.4).

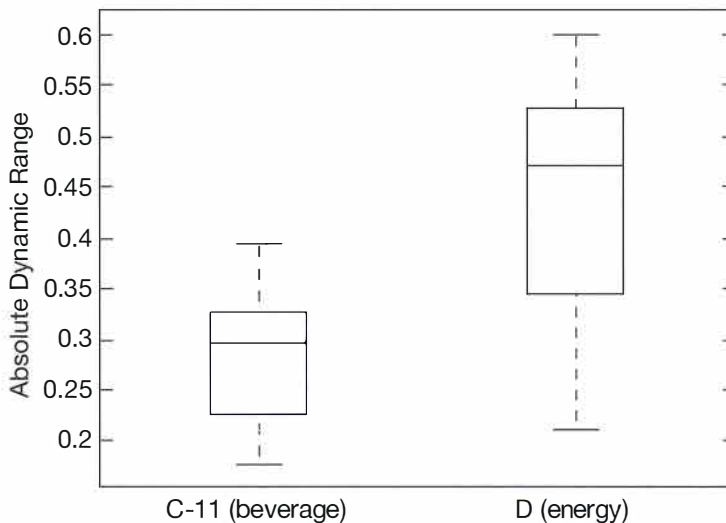
Es lassen sich beim Vergleich der Audiologos aus unterschiedlichen Branchen noch eine ganze Reihe weiterer Unterscheidungsmöglichkeiten anführen (die

**Abb. 6:**

Audiologo-Timbre-Space für Audiologos aus dem Einzelhandel (G) vs. denen aus der Elektronikbranche (C-26) mit den Dimensionen „Spectral Centroid“ (Hz), „Spectral Flux“ (euklidische Distanz) und „First Attack Time“ (Sekunden)

alle in der Matrix in Abbildung 3 zusammengefasst sind). Das Spannungsfeld, in dem sich die deutlichsten klanglichen Unterschiede am stärksten widerspiegeln, spannt sich zwischen den Audiologos aus der Filmbranche (J-59) und denen aus dem Einzelhandel (G) auf, während sich die paarweisen klanglichen Unterschiede der anderen Branchen zwischen diesen beiden Polen aufreihen lassen.

Zuletzt soll nicht unerwähnt bleiben, dass die Unterscheidbarkeit zwischen Audiologos beim Hörer natürlich nicht nur auf musikstrukturelle bzw. klangliche Merkmale zurückzuführen ist. Gerade bei simultaner Wahrnehmung unterschiedlicher Reizqualitäten (audiovisuellen Darbietung in Werbespots) wird stets versucht, eine kausale Verbindung herzustellen (Passung). Visuelle und akustische Eindrücke können sich dabei nicht nur ergänzen, sondern es können auch teils signifikante Varianzen bei auditiver (Audiologo) und visueller Darbietung (visuelles Markenlogo) beobachtet werden (Anzenbacher et al., 2016). Verlässliche Aussagen über eine tatsächliche Werbewirkung fordern daher sowohl den Einbezug der permanenten (intermodalen) Wechselwirkung verschiedener Reizqualitäten als auch des medialen Kontexts und bestehenden Markenwissens.

**Abb. 7:**

Vergleichender Boxplot über die Kanaldynamik (RMS-Energiewerte mit Abstand zwischen Maximum und 0.1-Perzentil) von Audiologos aus der Getränkeindustrie (G) im Vergleich zu denen aus der Energiebranche (D)

5 Fazit

Insgesamt lassen sich die untersuchten Audiologos anhand ihrer klanglichen und musikalischen Eigenschaften verschiedenen Branchen zuordnen. Sowohl bei Merkmalen wie dem spektralen Schwerpunkt, der Einschwingzeit, der verwendeten Instrumente, der Länge wie auch dem Auftreten und der Einsatzweise der menschlichen Stimme lassen sich signifikante Unterschiede innerhalb der verschiedenen Branchen feststellen. Gleichzeitig kommt es aber zu Einschränkungen in der Zuweisung zu Branchen durch anerkannte Klassifizierungsverfahren.

Es konnten darüber hinaus auch klangliche Unterschiede bezüglich der Herkunft der Audiologos festgestellt werden, jedoch könnten die Ursachen hierfür vielfältiger Natur sein, sodass an dieser Stelle noch keine Aussagen über eventuelle länder- oder kulturspezifische Klangmerkmale möglich sind. Klangliche Unterschiede hinsichtlich der Historizität lassen sich gleichwohl in der analysierten Population nicht ausschließen, aber aufgrund der sehr schwer zugänglichen Informationen nicht dokumentieren.

Für die Komposition von Audiologos bieten die vorliegenden Ergebnisse Orientierungspunkte, um sowohl auf der Grundlage der branchentypischen klanglichen Eigenschaften gezielt akustisch markengerechte bzw. -konforme Audiobotschaften zu erschaffen, als auch durch ungewohnte, klanglich untypische Strukturen einen neuen Grad an verkaufsfördernder Aufmerksamkeit zu erreichen. Die Erweiterung der Stichprobe mit Anpassung der Methodik sowie der zusätzliche Einbezug von Gedächtniseffekten sind Aufgaben zukünftiger Forschungsarbeiten.

Literatur

- Aichner, T. (2014). Country-of-origin marketing: A list of typical strategies with examples. *Journal of Brand Management*, 21 (1), 81–93. <http://doi.org/10.1057/bm.2013.24>
- Allan, D. (2007). Sound advertising. *Journal of Media Psychology*, 12 (3), 1–35.
- Alpert, M. I., Alpert, J. I. & Maltz, E. N. (2005). Purchase occasion influence on the role of music in advertising. *Journal of Business Research*, 58 (3), 369–376. [http://doi.org/10.1016/S0148-2963\(03\)00101-2](http://doi.org/10.1016/S0148-2963(03)00101-2)
- Anzenbacher, C. (2012). *Audiologos. Integrative Gestaltungsmaßnahmen vor dem Hintergrund der Musikpsychologie*. Baden-Baden: Nomos. <http://doi.org/10.5771/9783845270753-70>
- Anzenbacher, C. (2016). Das Audio Logo als akustische Visitenkarte einer Marke. In K. Bronner (Hrsg.), *Audio Branding* (S. 71–91). Baden-Baden: Nomos.
- Anzenbacher, C., Cziedik-Eysenberg, I., Reuter, C. & Oehler, M. (2015). Der Klang der Marken – Branchentypische psychoakustische Eigenschaften von Audiologos. In S. Becker (Hrsg.), *Fortschritte der Akustik – DAGA 2015* (S. 928–931). Berlin: Deutsche Gesellschaft für Akustik.
- Anzenbacher, C., Reuter, C. & Oehler, M. (2013). Sound quality vs. sound identity. The perceptibility of audio logos under everyday conditions of transmission and reception. *Audio Branding Academy Yearbook 2013/14* (S. 127–135). Baden-Baden: Nomos.
- Badri, M. A., Davis, D. L. & Davis, D. F. (1995). Decision support for global marketing strategies: the effect of country-of-origin on product evaluation. *Journal of Product & Brand Management*, 4 (5), 49–64. <http://doi.org/10.1108/10610429510103827>
- Bronner, K. & Hirt, R. (2009). *Audio-Branding: Brands, sound and communication*. Baden-Baden: Nomos. <http://doi.org/10.5771/9783845216935>
- Brooker, G. & Wheatley, J. (1994). Music and radio advertising: effects of tempo and placement. *Advances in Consumer Research*, 21, 286–290.
- Bruner, G. C. (1990). Music, Mood, and Marketing. *Journal of Marketing*, 54 (4), 94–104. <http://doi.org/10.2307/1251762>
- Cutieta, R. A. & Booth, G. D. (1996). The influence of metre, mode, interval type and contour in repeated melodic free recall. *Psychology of Music*, 24, 222–236. <http://doi.org/10.1177/0305735696242012>
- Dowling, W. J. (1978). Scale and contour – 2 components of a theory of memory for melodies. *Psychological Review*, 85 (4), 341–354. <http://doi.org/10.1037/0033-295X.85.4.341>
- Eiting, M. H. (1984). Perceptual similarities between musical motifs. *Music Perception*, 2 (1), 78–94. <http://doi.org/10.2307/40285283>
- Huron, D. (1996). The melodic arch in Western folksongs. *Computing in Musicology*, 10, 3–23.
- Jackson, D. (2004). *Sonic branding: An essential guide to the art and science of sonic branding*. Houndmills, UK: Palgrave Macmillan.
- Kellaris, J. J., Cox, A. D. & Cox, D. (1993). The effect of background music on ad processing: A contingency explanation. *Journal of Marketing*, 57, 114–125. <http://doi.org/10.2307/1252223>
- Kellaris, J. J. & Kent, R. (1991). Exploring tempo and modality effects, on consumer responses to music. *Advances in Consumer Research*, 18, 243–248.
- Lartillot, O. (2013). *MIRtoolbox 1.5 user's manual*. Finnish Centre of Excellence in Interdisciplinary Music Research. Retrieved from <https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox/MIRtoolbox1.5Guide>
- Lartillot, O. & Toivainen, P. (2007). A Matlab toolbox for musical feature extraction from audio. *Proceedings of the 10th International Conference on Digital Audio Ef-*

- fects (DAFx-07). Bordeaux, FRA. Retrieved from <http://dafx.labri.fr/main/papers/p237.pdf>
- Lartillot, O., Toivianen, P. & Eerola, T. (2008). A Matlab toolbox for music information retrieval. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme & R. Decker (Eds.), *Data analysis, machine learning and applications. Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation e.V., Albert-Ludwigs-Universität Freiburg, March 7–9, 2007* (pp. 261–268). Berlin: Springer.
- Lehmann, M. (2016). Die Stimme im Markenklang. In K. Bronner (Hrsg.), *Audio Branding* (S. 93–98). Baden-Baden: Nomos.
- MacInnis, D. & Park, C. (1991). The differential role of characteristics of music on high- and low-involvement consumers' processing of ads. *Journal of Consumer Research*, 18, 161–173. <http://doi.org/10.1086/209249>
- Müllensiefen, D. (2004). *Variabilität und Konstanz von Melodien in der Erinnerung: ein Beitrag zur musikpsychologischen Gedächtnisforschung*. Unveröffentlichte Dissertation, Universität Hamburg.
- North, A. C., Hargreaves, D. J., MacKenzie, L. & Law, R. (2004). The effects of musical and voice 'fit' on responses to adverts. *Journal of Applied Social Psychology*, 34, 1675–1708. <http://doi.org/10.1111/j.1559-1816.2004.tb02793.x>
- Peterson, R. A. & Jolibert, A. J. P. (1995). A meta-analysis of country-of-origin effects. *Journal of International Business Studies*, 26 (4), 883–900. <http://doi.org/10.1057/palgrave.jibs.8490824>
- Petty, R. E. & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of Persuasion. *Advances in Experimental Social Psychology*, 19, 123–205. [http://doi.org/10.1016/S0065-2601\(08\)60214-2](http://doi.org/10.1016/S0065-2601(08)60214-2)
- Plomp, R. & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America (JASA)*, 38 (4), 548–560. <http://doi.org/10.1121/1.1909741>
- Poth, L. (2001). Für erfolgreiche Händler sind die Mitarbeiter entscheidend. *Absatzwirtschaft*, 5, 56–62.
- Poth, L. & Poth, G. (1999). *Gabler Markenbegriffe von A-Z*. Wiesbaden: Gabler. <http://doi.org/10.1007/978-3-322-94419-1>
- Reuter, C. & Oehler, M. (2011). Psychoacoustics of chalkboard squeaking. *Journal of the Acoustical Society of America (JASA)*, 130 (4), 2545. <http://doi.org/10.1121/1.3655174>
- Reuter, C., Oehler, M. & Mühlhans, J. (2014). Physiological and acoustical correlates of unpleasant sounds. *Proceedings of the Joint Conference ICMPC13-APSCOM5*, August 4–8, 2014, Yonsei University, Seoul, Korea, 97. Retrieved from https://www.researchgate.net/publication/264557524_Physiological_and_acoustical_correlates_of_unpleasant_sounds
- Sethares, W. A. (1998). *Tuning, timbre, spectrum, scale*. London: Springer-Verlag. <http://doi.org/10.1007/978-1-4471-4177-8>
- Shen, Y. & Chen, T. (2006). When east meets west: The effect of cultural tone congruity in ad music and message on consumer ad memory and attitude. *International Journal of Advertising*, 25 (1), 51–70. <http://doi.org/10.1080/02650487.2006.11072951>
- Siddiq, S., Reuter, C., Czedik-Eysenberg, I. & Knauf, D. (2015). Vergleichende Untersuchungen zu Timbre Space Studien. In S. Becker (Hrsg.), *Fortschritte der Akustik – DAGA 2015* (S. 811–813). Berlin: Deutsche Gesellschaft für Akustik.
- Snyder, B. (2000). *Music and memory. An introduction*. Cambridge, MA: MIT Press.
- Steiner, P. (2014). *Sound Branding. Grundlagen der akustischen Markenführung*. Wiesbaden: Gabler.

- Tzanetakis, C. (2002). Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 10 (5), 293–302. <http://doi.org/10.1109/TSA.2002.800560>
- United Nations Statistics Division. (2014). *International standard industrial classification (ISIC) of all economic activities. Revision 4* (Statistical papers, Series M, No. 4/Rev. 4). Retrieved from http://unstats.un.org/unsd/publication/seriesM/seriesm_4rev4e.pdf
- Yalch, R. F. (1991). Memory in a jingle jungle: Music as a mnemonic device in communicating advertising slogans. *Journal of Applied Psychology*, 76 (2), 268–275. <http://doi.org/10.1037/0021-9010.76.2.268>