

# Diplomarbeit

## File-basierte Lautheitsmessungen bei Kinofilmen

---

### Perspektiven für eine optimierte Wiedergabelautstärke?

vorgelegt von Gregor Bonse, Matr.-Nr.: 4524  
betreut durch Prof. Martin Steyer

Mit freundlicher Unterstützung von:



## **Vorwort**

Kein Thema hat mich in meinem Studium so sehr beschäftigt, wie die Wahl der Abhörlautstärke während der Kinomischung und des Wiedergabelevels bei der Kinovorführung. Nicht selten habe ich eigene Mischungen revidiert, habe zu leise Wiedergaben im Voraus zu kompensieren versucht und bin durch lange Flure von Multiplex Kinos gelaufen, um den Vorführer zu finden, während im Saal bereits Rolle 2 anlief. Wenn ich die Debatte zur Kinolautstärke und die Diskussion um eine neue Aussteuerungs-Norm mit dieser Arbeit bereichern kann, dann ist mein Ziel bereits erreicht.

# Inhaltsverzeichnis

<b>Begriffserklärungen .....</b>	<b>5</b>
<b>Einleitung.....</b>	<b>6</b>
<b>1. Kinolautstärke - von der Produktion bis zum Konsum.....</b>	<b>10</b>
<b>1.1. Konsum .....</b>	<b>10</b>
1.1.1. Exkurs: TV .....	10
1.1.2. Film.....	12
1.1.3. Werbung und Trailer .....	13
<b>1.2. Auswertung .....</b>	<b>14</b>
1.2.1. Film.....	14
1.2.2. Werbung und Trailer .....	16
<b>1.3. Standardisierung und Regulierung.....</b>	<b>17</b>
1.3.1. 35mm-Lichtton .....	17
1.3.2. D-Cinema.....	18
1.3.3. Werbung und Trailer .....	19
<b>1.4 Produktion.....</b>	<b>19</b>
1.4.1. Film.....	19
1.4.2. Werbung und Trailer .....	23
<b>1.5 Zwischenbilanz.....</b>	<b>23</b>
<b>2. Lautheits-Messungen - Verfahren und Anwendungen .....</b>	<b>28</b>
<b>2.1. Objektive Lautstärke-Messmethoden .....</b>	<b>28</b>
<b>2.2. TASA Standard nach <math>L_{EQ(M)}</math>.....</b>	<b>30</b>
2.2.1. Initiative .....	30
2.2.2. Verfahren .....	31
2.2.3. Umsetzung.....	31
2.2.4. Erfolg .....	32
<b>2.3. AC-3 und Dialog-Normalisierung .....</b>	<b>33</b>
2.3.1. Verfahren .....	33
2.3.2. Dialogue Intelligence™ .....	34
2.3.3. Umsetzung.....	37
2.3.4. Erfolg .....	38

<b>2.4. EBU R 128 und ATSC A/85 nach ITU-R BS.1770 .....</b>	<b>40</b>
2.4.1. Differenzierung.....	40
2.4.2. Initiative.....	40
2.4.3. Verfahren.....	41
2.4.4. Umsetzung .....	44
2.4.5. Erfolg.....	46
<b>3. Praktische Lautheitsmessungen bei Kinofilmen .....</b>	<b>49</b>
<b>3.1. Auswahl der Algorithmen und zu erfassenden Kennwerte .....</b>	<b>49</b>
<b>3.2. Erstellung der Film-Datenbank .....</b>	<b>49</b>
<b>3.3. Subjektive Einschätzungen .....</b>	<b>52</b>
<b>3.4. Durchführung der Messungen .....</b>	<b>53</b>
3.4.1. NUGEN LMB .....	53
3.4.2. Dolby Media Meter 2 .....	54
<b>3.5. Auswertung .....</b>	<b>57</b>
3.5.1. Objektive Auswertung der Datenbank.....	57
3.5.2. Auswertung der Graphen.....	61
3.5.3. Beurteilung der Dialogue Intelligence Performance.....	63
3.5.4. Vergleich von Messergebnissen und subjektiver Einschätzung .....	66
<b>4. Konzeption einer Modellnorm.....</b>	<b>72</b>
<b>4.1. Optimierung der Dialoglevel-Bestimmung.....</b>	<b>72</b>
4.1.1. Implementierung eines relativen Gates.....	74
<b>4.2. Anpassung des Lautheits-Algorithmus .....</b>	<b>76</b>
<b>4.3. Nomenklatur .....</b>	<b>76</b>
<b>4.4. Maximale Lautheit.....</b>	<b>77</b>
<b>4.5. Festlegung der Zielwerte .....</b>	<b>79</b>
<b>4.6. Umsetzung .....</b>	<b>81</b>
<b>4.7. Auswirkung .....</b>	<b>84</b>
<b>Fazit .....</b>	<b>88</b>
<b>Literaturverzeichnis.....</b>	<b>57</b>
<b>Abbildungsverzeichnis .....</b>	<b>94</b>
<b>Tabellenverzeichnis .....</b>	<b>95</b>
<b>Selbstständigkeitserklärung .....</b>	<b>97</b>


## Begriffserklärungen

### *Lautstärke und Lautheit*

Während im Englischen allein das Wort „Loudness“ für den menschlichen Höreindruck von Schall steht, hat sich im Deutschen neben „Lautstärke“ noch „Lautheit“ etabliert. Eingeführt wurde dieser Begriff ursprünglich durch Stanley S. Stevens, um dem psychoakustischen logarithmischen Vergleichsmaß Lautstärkepegel in phon eine proportionale Abbildung des menschlichen Lautstärkeempfindens gegenüberzustellen, die Lautheit in sone. Sowohl Lautstärkepegel als auch Lautheit sind also besetzte Größen der Lautstärke. Trotzdem wird auch die in der EBU R 128 definierte „Loudness“ als „Lautheit“ übersetzt, da sich der Begriff im Rundfunkbereich bereits umgangssprachlich durchgesetzt hat. Um Missverständnisse zu vermeiden, soll in dieser Arbeit also dann von „Lautstärke“ die Rede sein, wenn im klassischen Wortsinn der menschliche Höreindruck gemeint ist, und von „Lautheit“ nur dann, wenn es explizit um Größen wie Programmlautheit, Lautheitsbereich etc. geht, wie sie in der R 128 und bestimmten Algorithmen definiert sind.<sup>1</sup>

Sämtliche Berufsbezeichnungen sind geschlechtsneutral zu verstehen.

---

<sup>1</sup> siehe auch  EBU: Empfehlung R 128, S. 3

## Einleitung

*„I mix on 7 in my calibrated room, make the movie sound great and the director and producers happy, cash the check, (...) then, I move on.“<sup>1</sup>*

Mit diesem Statement erklärt ein anonymer Mischtonmeister seine Sichtweise auf das viel diskutierte Problem der Kino-Lautstärke. Er mische auf [Dolby Level] 7 in seinem kalibrierten Studio, Sorge dafür, dass der Film großartig klinge und Regisseur sowie Produzenten zufrieden seien, löse den Scheck ein und widme sich dann Neuem.

Damit ist eigentlich alles gesagt. Die grundlegenden Aufgaben des Mischtonmeisters sind abgedeckt und er hält die Vorschrift von Dolby ein, welche die Mischungsbedingungen für Kinofilme mit Dolby SRD definiert. Tatsächlich ist die Aussage aber mit einigem Trotz verbunden, Trotz und Resignation gegenüber allem, was nach der Fertigstellung des Films passiert und sich seines Einflusses entzieht. Seine Einstellung habe sich diesbezüglich über die Jahre stark verändert.<sup>2</sup> Der anfängliche Idealismus für das eigene kreative Werk und das Selbstverständnis, auch allen Zuschauern nach ihm den gleichen Genuss zu ermöglichen, die gleiche Lautstärke, den Film so wie vom Regisseur während der Mischung gehört, weicht mit der Zeit einer Art Selbstschutz und der Besinnung auf das Wesentliche:

*„(...) accept the fact that once the mix has left your calibrated dubbing stage, you don't have any control over it anymore. (...) If you don't, you'll grow old and wear out very quickly. After all, it's a great job and a lot of fun.“<sup>3</sup>*

Mit dieser Ansicht steht der Mischtonmeister nicht alleine da, sie bildet aber im Kreis der Tonschaffenden die moderate Minderheit. Andere wollen sich noch nicht mit der Tatsache zufrieden geben, dass ihre Arbeit durch leise Pegel abgewertet wird, sind frustriert und äußern sich auch entsprechend:

*„I see this as a serious threat to our industry! Why pay the price of a Dolby approved mix stage if the mix and levels are no more likely to be correct than some guy with PTLE [Pro Tools Lite Edition] in his garage studio? For that matter why pay for a professional mixer in the first place? Just give the picture editor a cheap home cinema surround sound system and a few extra bucks to do it! A dynamic, exciting, high quality mix played back on a full range calibrated cinema system is one of the main elements which separates the cinema experience from watching a BluRay on a big screen at home.“<sup>4</sup>*

Was also zunächst nur ein wunder Punkt im Ton Department zu sein scheint, ein Dorn im Auge desjenigen, der sein Schaffen angefochten sieht, bekommt hier eine marktwirtschaftliche Komponente. Eine unbefriedigende Audiowiedergabe darf konkret mit einer Abwanderung vom Kinokonsum in Verbindung gebracht werden.

Es beschwerten sich nämlich nicht nur Tonmeister. Kritisch diskutiert wird auf Premieren und Festivals, bei kommerziellen Vorführungen, in Internetforen, unter Produzenten, Regisseuren, Kinobetreibern, Filmvorführern und, am allerwichtigsten, unter Konsumenten. In diesem Kreis gibt es zwar sehr wohl Enthusiasten, die es gerne lauter im Kino hätten und die deshalb ihr Heimkino bevorzugen<sup>5</sup>; die Mehrheit der Kinogänger klagt jedoch über zu laute Pegel. Das wirkt zunächst wie eine unüberbrückbare Differenz zwischen Filmschaffenden und Publikum.

Kompensationsversuche auf beiden Seiten führen durch Wechselwirkung schließlich zum gegenwärtigen Teufelskreis: Auf Seiten der Produktion (Produzenten, Regisseure, Mischtonmeister) wird lauter gemischt und auf Seiten der Auswertung (Kinobetreiber, Filmvorführer) wird leiser projiziert. Je lauter gemischt wird, desto leiser wird projiziert und je leiser projiziert wird, desto lauter wird gemischt. Fügt man die weiterhin nach Standard arbeitenden Parteien in die Gleichung mit ein, also die auf Level 7 arbeitenden Tonmeister und Vorführer, so erhält man einen Zustand von „free floating playback levels“,<sup>6</sup> wie ihn der deutsche Mischtonmeister Matthias Lempert nennt. Niemand kann vorhersehen, auf welcher Lautstärke ein Film am Ende sein Publikum erreicht; die Regler schwanken sowohl in der Mischung als auch bei der Auswertung irgendwo zwischen den Extremen. Genaue Werte sind schwer festzumachen, da es keinerlei Erhebungen zum Thema gibt. Allgemein ist aber festzustellen, dass zwischen 3 und 7 alles möglich ist, wengleich die Wahrscheinlichkeit zum Mittelfeld hin zunimmt.<sup>7,8,9</sup> Das kann fatale Ausmaße annehmen. So wird ein nach Vorschrift bei Level 7 gemischtes Dialog-Drama auf Level 4 projiziert, viel zu leise sein, möglicherweise nicht mehr genügend Sprachverständlichkeit bieten, geschweige denn, dramaturgisch wichtige Tongestaltung transportieren. Ein Action-Thriller hingegen, der auf Level 5 oder weniger gemischt wurde um konkurrenzfähig zu sein, wird, bei 7 projiziert, viel zu laut sein und vielleicht in dieser Lautstärke sogar gesundheitsschädliche Ausmaße annehmen.

Von einem „Loudness War“, also einem Krieg um die Lautheit, muss spätestens dann die Rede sein, wenn man Filmtrailer und Werbungen einbezieht, die mit ihren kategorisch lauten Mischungen die genannte Überkompensation am Kinoprozessor weiter provozieren.

In dieser Arbeit möchte ich untersuchen, inwiefern das Problem der Wiedergabelautstärke im Kino durch File-basierte Lautheitsmessungen entschärft werden könnte. Die zentrale Frage dabei lautet: Wenn eine Mischung bewusst zu

laut erstellt wurde, bei Abhörlautstärken unter Referenzlevel, kann diese Abweichung von der Norm dann technisch nachempfunden werden? Sind moderne Lautheits-Algorithmen in der Lage, diesen Fehler auszumachen und wenn ja, bieten sie dabei eine ausreichende Präzision? Inwiefern beeinflussen Genres und somit Mischungskonzepte die Lautheits-Messungen? Letztendlich muss untersucht werden, ob mit den heute zur Verfügung stehenden Mitteln eine sinnvolle Modellnorm für das digitale Kino der Zukunft formuliert werden könnte und welche Auswirkungen eine solche auf Gestaltung und Auswertung hätte.

Unabhängig davon, ob das Inkrafttreten einer solchen Vorschrift realistisch oder überhaupt wünschenswert ist, halte ich diese Untersuchung für substantiell wichtig, auch wenn sie nur die hypothetischen Optionen abwägt. Spätestens seit der R 128 Empfehlung der EBU und der damit verbundenen, in Deutschland seit August 2012 wirksamen Lautheits-Normalisierung im Fernsehen, sollten auch die Möglichkeiten für die Kino-Distribution neu ausgewertet werden. Und wengleich eine gewisse Stagnation in der Debatte zu verzeichnen ist und einflussreiche Unternehmen, Verbände und Institutionen sich mit dem Status Quo abgefunden haben, so gibt es doch weiterhin zahlreiche, unlängst sogar prominente Stimmen aus der Branche, die eine solche Regelung favorisieren.










Randy Thom (Mischtonmeister zu Forest Gump, The Incredibles, Harry Potter 2 & 4):

*„I think the solution will require that the Producers', Directors', and Theater Owners' trade organizations come to an agreement about guidelines for the average loudness and peak loudness for movies.“<sup>10</sup>*

Matthias Lempert (Mischtonmeister zu Cloud Atlas, Pina, Das Parfum):

*„We have to face it - cinema needs sth. like EBU-R128 loudness regulations to get rid of these level inconsistencies between the mixing stages and the cinemas.“<sup>11</sup>*



- 
- <sup>1</sup>  DUC: Dolby 7. Is everybody mixing at this level? (mr.armadillo, 05.19.2011, 12:39)
- <sup>2</sup> vgl. ebenda
- <sup>3</sup> ebenda: „Akzeptiert, dass ihr keine Kontrolle mehr über die Mischung habt, wenn sie erstmal euer kalibriertes Studio verlassen hat. Wenn ihr das nicht tut, werdet ihr altern und euch sehr schnell abnutzen. Letzten Endes ist es ein großartiger Beruf und eine Menge Spaß.“
- <sup>4</sup> vgl.  DUC: Cinema Chains - sadness at poor levels (Greg Malcangi, 01.05.12, 08:50)  
„Ich sehe das als ernsthafte Bedrohung für unsere Industrie! Warum sollte man Geld für eine Dolby lizenzierte Mischung ausgeben, wenn die Wahrscheinlichkeit, dass Mischung und Pegel korrekt sind, ebenso gering sind wie bei Jemandem, der mit Pro Tools LE in seiner Garage mischt? Warum überhaupt einen professionellen Mischtonmeister engagieren? Man könnte genauso gut dem Cutter ein Heimkino Surround System und ein paar Dollar mehr geben und ihn den Job machen lassen. Eine dynamische, spannende, hochqualitative Mischung, wiedergegeben auf einer kalibrierten Full-Range Kinoanlage ist eines der Hauptelemente, die das Kinoerlebnis vom Heimkinoerlebnis mit Bluray und großem Bildschirm trennt.“
- <sup>5</sup> vgl.  Cinefacts: Blu-ray vs. Kino? (Calypso, 05.02.10, 20:35)  
„Ich finde den Ton im Kino meistens viel zu leise, da gefällt mir der Sound zuhause deutlich besser.“
- <sup>6</sup>  DUC: Dolby 7. Is everybody mixing at this level? (matt-o-, 14.05.11, 06:05)
- <sup>7</sup>  BVFT: Abhörlautstärke bei Trailermischungen (Stefan Rüdell, 06.02.09, 19:38)  
„Gerade heute wieder eine Diskussion mit einem Kunden, der schon beim Reinkommen den Wunsch äußert, ‚doch bitte auf 3 zu mischen‘, weil es hinterher im Kino sonst wieder ‚zu leise‘ ist.“
- <sup>8</sup>  DUC: DCP cinema volume discrepancy (Brent\_Sydney, 22.03.12, 18:32)  
„(...) 99% of stuff seems to be played at 5. 1% at 6. nothing at 7.“
- <sup>9</sup>  Kevin Markwick, Cinema Owners Blog - Can you hear me at the back?  
„(...) when I turn the volume to 7, it's exactly at the level the director wants you to hear the film.“
- <sup>10</sup>  Thom, Randy (Sound-Article-List): The future of hollywood sound... (09.03.09, 01:56)  
„Ich denke die Lösung bedarf einer Einigung der Gewerkschaften von Produzenten, Regisseuren und Kinobesitzern über Richtlinien für durchschnittliche und maximale Lautheit von Filmen.“
- <sup>11</sup>  DUC: Cinema Chains - sadness at poor levels (matt-o-, 05.05.12, 09:09)  
„Wir müssen es akzeptieren - Kino braucht so etwas wie die EBU-R128 Lautheitsvorschrift um diese Pegel Unkontinuitäten zwischen Mischungsstudios und Kinos zu beseitigen.“

# 1. Kinolautstärke - von der Produktion bis zum Konsum

Um verbesserte Lautstärke-Bedingungen zu schaffen, müssen zunächst die einzelnen Interessen und Verhaltensmuster genauer untersucht werden. Ein ideales System aus Produktion, Standardisierung sowie Regulierung, Auswertung und Konsum muss jede der Interessen berücksichtigen und Schnittmengen ausnutzen sowie gezielte Kompromisse eingehen. Eine Priorisierung der einzelnen Parteien stellt sich dabei als Gratwanderung von philosophischer Qualität heraus, da sie zwischen Kunst und Kommerz gewichtet: Bleibt die Kontrolle über Lautstärke als Gestaltungsmittel allein bei den Filmschaffenden, wird der Kunstcharakter des Filmes gestärkt. Entscheidet allein die Kinoauswertung über die Wiedergabe, wird das Werk zur Ware. Die Interessen und Verhaltensmuster werden vom Konsum bis zur Produktion zurückverfolgt und anschließend untersucht, wie eine Synthese auszusehen hätte.

## 1.1. Konsum

### 1.1.1. Exkurs: TV

Der Frage „Was befriedigt Hörer?“ wurde bereits 2003 von Jeffrey C. Riedmiller et al. der Firma Dolby Laboratories nachgegangen, welche die Ergebnisse in einem AES Convention Paper publik machten.<sup>1</sup> In

der Studie sollte untersucht werden, welches Maß an Lautstärke-Unterschieden innerhalb und zwischen Programmen von Fernsehzuschauern akzeptiert wird. Das Ergebnis wurde als „Komfort-Zone“ bezeichnet und ist in Abbildung 1 dargestellt. Die persönlich bevorzugte Fernsehlautstärke, nennen wir sie Ideallautstärke (hier 0dB), wurde von den Testpersonen einmalig eingestellt und anschließend untersucht, welches Maß an Akzeptanz oder

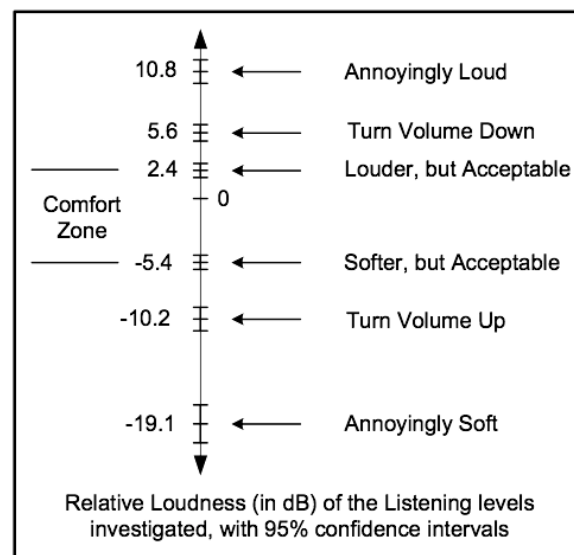


Abbildung 1: Komfort-Zone TV Zuschauer

Ablehnung eine Pegeländerung nach oben und unten hervorruft. Als Testsignal diente hierbei Sprache. Interessant ist, dass die Skala asymmetrisch verläuft: Während ein Klangbeispiel bereits wenige dB über der Ideallautstärke zunehmend

störend wird und bei durchschnittlich 5,6dB schließlich der Griff zur Fernbedienung folgt, darf es bis zu etwa 10,2 dB leiser werden, bevor der Konsument den Fernseher lauter stellt. Zwei Einschränkungen sind hierbei zu beachten: Die Tests fanden unter Bedingungen statt, die etwa dem eines Wohnzimmers zu ruhiger Tageszeit entsprechen; steigende Störgeräusche würden die unteren Akzeptanzgrenzen negativ beeinflussen. Außerdem können die Tests an Sprache nicht einfach auf Fernsehprogramme per se übertragen werden. Es kann nicht davon ausgegangen werden, dass bei lauten Geräuschen oder lauten Musiksequenzen in einem Film gleich leiser geschaltet wird, wenn diese 6 dB lauter als Sprachpassagen sind, also über dem Wert "turn volume down" liegen. Dolby geht davon aus, dass nicht-sprachliche Elemente - sofern sie ausgewogen um den Dialog herum gemischt wurden - Zuschauer nicht irritieren, solange die Sprache in der Komfort-Zone liegt. Zuvor wurde nämlich untersucht, welche Elemente ein Hörer seiner Lautstärke-Einschätzung zugrunde legt, speziell bei welchen Beispielklängen Testpersonen besonders sicher Lautstärken in Relation zueinander setzen können. Sprache setzte sich dabei klar von Musik, anderen

Geräuschen und Testsignalen, wie Rauschspektren und Tönen, ab. Abbildung 2 stellt die Übereinstimmung von 21 Hörern dar, die jeweils ein Klangbeispiel, das Sprache enthält, und eines, das lediglich aus Schritt-Geräuschen besteht, in Relation zu einem Referenzsignal setzen sollten. Diese Ergebnisse sollten nicht so verstanden werden, dass etwa Testperson A ein Programm aus

Geräuschen 12 dB leiser hören möchte als Testperson B. Vielmehr sind ihre Angaben beliebiger als in Bezug auf Sprache. Ein Zuschauer weiß so lange nicht, ob er die richtige Lautstärke eingestellt hat, bis das erste Sprachelement einer Sendung kommt, weil Geräusche wie Schritte einen viel größeren Toleranzbereich haben.

Riedmiller et al. schließen also mit dem logischen Fazit, dass eine zukünftige Anpassung verschiedener Programmanteile im Fernsehen auf deren Dialoglevel beruhen sollte und dieses durch Lautheitsmessungen von Dialogpassagen ermittelt

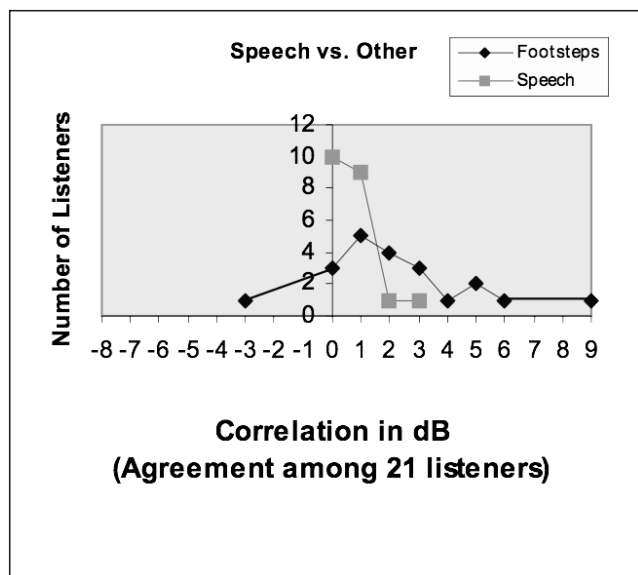


Abbildung 2: Übereinstimmung bei Lautstärke-Wahrnehmung von Sprache und Geräuschen

werden müsse. Der Zuschauer könne im Folgenden selbst entscheiden, wo seine Komfort-Zone und somit seine Ideallautstärke liegt. Dolby stellt fest, dass diese bei Fernsehzuschauern in einer Wohnzimmeratmosphäre genau in jenem Lautstärkebereich liegt, in dem alltägliche Konversation stattfindet (55-66 dB<sub>A</sub>), und beruft sich dabei auf Studien von Benjamin<sup>2</sup> und Pearsons et al.<sup>3</sup>

### **1.1.2. Film**

Überträgt man diese Erkenntnisse nun auf einen Kinozuschauer, muss zunächst festgestellt werden, dass dieser seine Abhörlautstärke nicht wählen kann. Er hätte, selbst wenn er wollte, nicht die Möglichkeit, diese frei und individuell zu korrigieren, wenn seine Komfort-Zone über- oder unterschritten wird. Das Prinzip Kino stellt bereits in sich einen Kompromiss dar und folgt zwangsläufig den Regeln der Gruppendynamik.

Die Einflussnahme des Publikums beschränkt sich auf die Beschwerde beim Kinopersonal und hier vermitteln uns die Untersuchungen zur Komfort Zone ein besseres Verständnis für die gegenwärtige Situation: Zweifelsohne liegt die Schwelle, vom Kinositz aufzustehen, sich ggf. an weiteren Zuschauern vorbei zu bemühen und schließlich den Vorführer zu bitten, die Lautstärke zu ändern, höher, als die nach der Fernbedienung zu greifen. So läge beispielsweise der Wert, bei dem ein Zuschauer sich beim Vorführer über zu leise Pegel beschweren würde, vermutlich noch unter dem TV-bezogenen „Turn Volume Up“-Wert von -10,2dB, womöglich eher hin zu „Annoyingly Soft“, „Ärgerlich Leise“ bei -19,1dB (Abbildung 1). Wenn davon ausgegangen wird, dass die durchschnittliche Ideallautstärke der Zuschauer bei Dolby Level 7 liegt, was noch sehr optimistisch ist, entspricht letzterer Wert etwa der Reglerstellung 3 am Kinoprozessor<sup>4</sup>, die nur höchst selten vorkommt.

Erschwerend hinzu kommen allerdings kinotypische Störgeräusche von Popcorn- und Chipsgenuss sowie Gespräche im Publikum, die andere Ausmaße erreichen als in einer Wohnzimmeratmosphäre und die in der Dolby Studie ohnehin explizit nicht berücksichtigt wurden.

Fasst man diese Faktoren zusammen, wird deutlich, dass die Wahrscheinlichkeit einer Beschwerde eines durchschnittlichen Zuschauers über zu geringe Lautstärke im Kino verschwindend gering ist!

Die Wahrscheinlichkeit einer Beschwerde über zu laute Pegel ist höher. Als erster wird sich derjenige Zuschauer beschweren, dessen obere Akzeptanzgrenze überschritten wird. Ob diese nun bei +5,6dB „Turn Volume Down“ oder bei +10,8dB

„Annoyingly loud“ liegt, sei zunächst dahingestellt. Die Wiedergabelautstärke wird also nach seiner Beschwerde gesenkt und eine Gegenkorrektur kann so schnell nicht erwartet werden, da die Komfort-Zonen asymmetrisch verlaufen. Jedoch gibt es natürlich auch das andere Extrem eines Zuschauers, dessen Komfortzone besonders hoch liegt und der sich ab einem bestimmten, für ihn zu leisen Wiedergabelevel beschweren wird. Folgend muss sich, statistisch betrachtet, in einem System, dessen einzig wirkende Kraft die des sich beschwerenden Konsumenten ist, die Wiedergabelautstärke auf einen Wert einpendeln, der unterhalb der mittleren Ideallautstärke der Gruppe liegt. Um diesen Fehler genau zu beziffern, müsste untersucht werden, wo genau die „Beschwerde“-Werte in Relation zur Komfort-Zone bei Kinozuschauern liegen. Gingen wir davon aus, dass diese in etwa „Annoyingly Loud“ und „Annoyingly Soft“ entsprechen und dass sich die Wiedergabelautstärke dann auf lange Sicht in der Mitte zwischen den „Beschwerde“-Werten etabliert, dann wäre dieser Wert -4,15 dB unter der mittleren Ideallautstärke des Publikums<sup>5</sup>. Gingen wir nochmals von der optimistischen Annahme aus, letztere läge bei Dolby Level 7, dann wäre in einem solchen System Level 5,7 am Kinoprozessor realistisch; ein Bereich, der bei kommerziellen Screenings nur allzu geläufig ist. Den Konsens der Allgemeinheit spiegelt das Level jedoch nicht wieder, es resultiert lediglich aus der Tatsache, dass die jeweils extremen Wiedergabe-Vorstellungen bzw. die empfindlichsten Zuschauer beim Beschwerde-Prinzip dominieren.

### ***1.1.3. Werbung und Trailer***

Spezifiziert man die Interessen der Konsumenten nun auf Kinowerbung und Trailer, dann könnte zwar berechtigterweise angemerkt werden, dass diese von einigen am liebsten gar nicht geschaut würden und von anderen wiederum besonders gerne, es läge aber kein triftiger Grund vor, diese lauter oder leiser als den Hauptfilm konsumieren zu wollen. Diese Annahme spiegelt sich auch in der R 128 Empfehlung der EBU für ein Hörfunk- und Fernsehpublikum wieder.<sup>6</sup> Werbungen unterliegen hier den gleichen Lautheits-Vorgaben wie andere Programme, wengleich diese für kurze Spots aufgrund geringerer Integrations-Dauern Sender-seitig teilweise angepasst wurden.<sup>7</sup>

## 1.2. Auswertung

### 1.2.1. Film

Relevant für die Lautstärke im Kino bei der Auswertung sind die Interessen und das Verhalten der Kinobetreiber sowie Filmvorführer (Jene der Verleiher werden zur Produktion gezählt). Gelegentlich wird ein Film heutzutage vorweg durch das Kinopersonal in Gänze angesehen, um eine passende Initialpegelung für alle folgenden Vorführungen desselben Filmes zu finden. Andere Kinos stellen die Lautstärke jedes Mal nach dem Starten der Projektion und Testhören im Kinosaal neu ein; wohlgemerkt ist dann ein kurzer Ausschnitt ausschlaggebend für die Vorführung des gesamten Films. Ein weiteres, häufiges Szenario ist, dass der Dolby Regler kaum verstellt wird und letztlich auf einem Wert gehandhabt wird, der für den gesamten Content die wenigsten Probleme bereitet. Auch wird berichtet, dass bestimmte Kino-Ketten Obergrenzen am Kinoprozessor vorschreiben (z.B. kategorisch mindestens 6dB unter Referenzlevel, also zwischen Level 5 und 5,5 am Kinoprozessor).<sup>8</sup> Folgende Interessen der Kinobetreiber und Vorführer wirken auf die Wiedergabelautstärke ein:

1. Kundenzufriedenheit: Die Anzahl der unzufriedenen Kinozuschauer muss möglichst gering gehalten werden um Rückerstattungen zu vermeiden und einen erneuten Kinobesuch zu bewirken, weshalb auf Beschwerden nach aller Regel reagiert wird. Die Auswirkungen wurden in Kapitel 1.1.2. ausgeführt. Dabei wird auch das Kinopersonal mit der Zeit die Erfahrung machen, dass Beschwerden wegen zu hoher Lautstärke wahrscheinlicher sind als solche wegen zu geringer Lautstärke, wodurch es nahe liegt von vorne herein eher zu leise als zu laut zu projizieren.
2. Schutz des Wiedergabesystems: Als Grund, warum ein Film nicht lauter wiedergegeben werden könne, wird gelegentlich angegeben, das Wiedergabesystem müsse vor zu laut gemischten Filmen geschützt werden. Sicherlich wird das auch oft als Vorwand benutzt, wenn in Wirklichkeit z.B. Weisungen von oben befolgt werden, die eine Änderung des Levels untersagen. Einige Kinos sind aber in der Tat noch heute mit Kinoanlagen ausgestattet, die auf analoge Lichttonpegel ausgerichtet sind. Digitaler Filmtone kann auf solchen Systemen also speziell im Bereich von -14dBFS (100% Vollaussteuerung Lichtton) bis 0dBFS zu Verzerrungen führen und ggf.

Verstärkern sowie Lautsprechern schaden, wenn keine Leistungsreserve (safety margin) existiert. Dolby verlangte 1994 während der Übergangszeit von Dolby SR zu SRD (Dolby Digital) ein Wiedergabesystem, das lediglich in der Lage sei „Dolby SR Filme bei 100% Vollaussteuerung ohne Probleme über den gesamten Frequenzbereich wiederzugeben“. Dieses sei auch dann noch tragbar, wenn Kinos bereits Dolby Digital Filme zeigten; empfohlen sei jedoch dann eine leistungsfähigere Anlage, die  $103\text{dB}_C$  pro Frontkanal bewerkstelligen könne.<sup>9</sup> Heute werden  $105\text{dB}_{\text{SPL}}$  vorgeschrieben.<sup>10</sup>

3. Transmissionsschutz: Laute Szenen eines Filmes, speziell mit hohem LFE-Kanal-Anteil, sollten idealerweise im Kino nebenan nicht gehört werden. Genau das ist aber der Fall, wenn die Kinos schlecht schallgedämmt sind. Auch hierzu macht Dolby bestimmte Vorgaben: Die Summe aus STC (Sound Transmission Class, ein Maß der Schalldämmung zwischen zwei Räumen) und NC (Noise Criteria, ein Maß des Störpegels, der seinerseits Transmissionsschall verdecken kann) solle größer gleich 95 sein. Kinobetreiber, die Filme unter schlechteren Bedingungen zeigen, welche die Dolby Standards nicht erfüllen, verringern deshalb häufig die Wiedergabelautstärke bzw. den LFE-Kanal-Anteil.

4. Lizenz-Konformität:

- Dolby/DCI/SDDS/Datasat: Es ist ein weit verbreitetes Gerücht, Kinos seien Dolby gegenüber verpflichtet Filme auf Level 7 zu projizieren. Im Handbuch zum aktuellen Kinoprozessor CP750 steht Folgendes:

*„As with previous generations of Dolby cinema processors, a fader setting of 7.0 (0 dB) is the nominal correct operating level. 7.0 matches the level used during production of the motion picture.“<sup>11</sup>*

Die Vorgaben von Dolby haben Empfehlungs-Status. Ein Kino, das sich nicht an diese Empfehlungen hält, kann seine Dolby-Lizenz nicht verlieren, da Dolby-Lizenzen für Kinos gar nicht existieren. Ebenso wenig sind Konsequenzen von der DCI (Digital Cinema Initiatives, Herausgeber der DCP Norm) oder alternativen 35mm Digitaltonverfahren wie DTS (heute Datasat) und SDDS möglich. Auch letztere vertreiben Kinoprozessoren mit Referenzlevel bei 7,0 und sehen identische Einmessungen wie Dolby vor.<sup>12,13</sup>

- THX: Anders sieht es aus, wenn Kinos eine THX-Lizenz erwerben. Dieses Qualitätssiegel wird gegen Lizenz-Gebühren ausschließlich an Kinos vergeben, die strenge Kriterien zur Ton- und Bildwiedergabe erfüllen. Verstöße gegen diese Auflagen können selbstverständlich zum Verlust der Lizenz führen; jedoch schreibt auch THX kein definiertes Wiedergabe-Level vor. Auf Anfrage positioniert sich das Unternehmen folgendermaßen:

*„We do not require THX Certified Cinemas to reproduce movie soundtracks at ‚Dolby CP level 7.‘ The cinema management has full control of the volume in auditoriums and can adjust the level accordingly with the audience’s feedback.“<sup>14</sup>*

### **1.2.2. Werbung und Trailer**

Kinos finanzieren sich nicht nur durch den Verkauf von Eintrittskarten, sondern auch durch Werbungen und Trailer vor dem Hauptfilm. Interesse der Betreiber kann es also nicht nur sein, den Zuschauern gerecht zu werden, sondern sie müssen auch die Interessen ihrer Werbekunden berücksichtigen. Dennoch werden Werbungen und Trailer in der Regel weit unter Referenzlautstärke projiziert. Ironischerweise geschieht durch die Anlieferung lauter Spots nicht selten eine leisere Wiedergabe als beim Hauptfilm. Der bloße Gang zum Regler initiiert offenbar einen Gedanken, der für den Werbe-Effekt fatal ist: Wenn die Lautstärke für Werbespots ohnehin vermindert werden muss und diese das Publikum weniger interessieren als der Hauptfilm, wieso sollte dann diese Minderung nicht gleich großzügig sein? Unter dem Gesichtspunkt, dass ständiges Regeln der Wiedergabelautstärke Personal beansprucht und Kinos nicht selten versuchen, einen Mittelwert am Kinoprozessor für den gesamten Programminhalt zu finden, müssen Werbungen und Trailer zwangsläufig zu den Hauptgründen für zu leise Wiedergaben von Hauptfilmen gezählt werden. Nämlich genau dann, wenn das Level nach dem Werbe- und Trailerblock nicht wieder angepasst wird. Auch hier lässt sich ein ähnlicher Effekt wie bereits beim Konsum feststellen und zwar, dass zu laute Inhalte die Wahl des Dolby Levels unverhältnismäßig stärker beeinflussen als zu leise und gerade beim „Ein Level für Alles“ - Prinzip letztlich immer leisere Wiedergaben resultieren als das bei einer gleichmäßigen Mittelung der Fall wäre. Interesse der Kinobetreiber und Filmvorführer muss die Anlieferung von Spots sein, welche auf der gleichen Saalregler-Stellung projiziert werden können wie Hauptfilme.



## 1.3. Standardisierung und Regulierung

### 1.3.1. 35mm-Lichtton

Mitte der 1970er Jahre konzipierte Dolby Laboratories eine Empfehlung zur Kalibrierung von Monitor-Pegeln um einheitliche Wiedergabe-Bedingungen zu gewährleisten. Demnach sollte Rosa Rauschen bei 50% Lichtton-Aussteuerung (Dolby Level) auf Referenzlevel 7 am Kinoprozessor  $85 \text{ dB}_C$  pro Frontkanal und Mono-Surroundkanal entsprechen. Beim später eingeführten Dolby SR wurde dieses Dolby Level dann zur Steigerung des Headrooms auf  $88 \text{ dB}_C$  angehoben. Bei Dolby Digital wiederum sollte das Dolby Level laut Standard  $-20 \text{ dB}_{FS}$  entsprechen und seinerseits  $85 \text{ dB}_C$  pro Frontkanal sowie  $82 \text{ dB}_C$  pro Surround-Kanal erzeugen. Der zusätzliche LFE-Kanal soll bei Dolby Level auf einem Frequenz-Analyzer bis 125 Hz 10dB lauter sein als ein Fullrange-Lautsprecher. Dolby machte außerdem die Vorgabe, dass die Wiedergabe-Systeme auf die Kurve X eingemessen werden sollten, welche tiefe Frequenzen unterhalb von 63 Hz sowie hohe Frequenzen oberhalb von 2 kHz je nach Saalgröße absenkt. Während diese Abhörbedingungen für Kinos Empfehlungsstatus haben, sind sie für Dolby lizenzierte Studios, in denen Printmaster für Dolby SRD erstellt werden, Voraussetzung.

Somit nimmt Dolby einzig und allein durch die Vorgabe der Abhörbedingungen, inklusive des Wiedergabelevels bei Stellung 7, Einfluss auf die Mischung von Kinofilmen. Auch auf Nachfrage bestätigt das Unternehmen, dass Richtlinien zum Dialoglevel und Maximalpegel nie existierten, sondern ein Dialog-Zielwert von  $-20 \text{ dB}_{FS}$  sich mit der Zeit automatisch aus den Abhörbedingungen in der Praxis ergeben hat.<sup>15</sup> Es steht dem Mischtonmeister frei, wie er seine Dialoge und die restlichen Ton-Elemente mischt. In der Praxis wird gegenwärtig häufig entgegen der Vorgaben von Dolby unterhalb des Referenzpegels gemischt und erst für das Printmastering wieder lauter gestellt. Der Dolby Consultant greift in der Regel auch dann nicht ein, wenn Filme beim Printmastering zu laut klingen, sondern stellt lediglich sicher, dass die Abhörbedingungen stimmen. Innerhalb dieser Rahmenbedingungen besteht künstlerische Freiheit. Eine eigentliche Regulierung findet also nicht statt, da hierfür keine Grundlage besteht.

Das Verfahren von Dolby wurde später von der Society of Motion Picture & Television Engineers (SMPTE) im Dokument RP200 zur Einmessung von Kinos standardisiert. Auch hier wird lediglich darauf verwiesen, dass der Kinoprozessor

für ein normales Film-Programm im Kino auf dem kalibrierten Referenzlevel gehalten werden sollte.<sup>16</sup>

Die 1993 eingeführten digitalen Lichtton-Formate DTS (heute DATASAT) und SDDS, die in den USA noch immer verbreitet sind, passten ihre Vorgaben zur Kino- sowie Studioeinmessung ebenfalls dem SMPTE Standard an.

### **1.3.2. D-Cinema**

Zur Standardisierung eines vollständig digitalisierten Kinos wurde im Jahr 2002 von den Filmstudios Disney, Fox, Paramount, Sony Pictures Entertainment, Universal und Warner Brothers die Digital Cinema Initiatives (DCI) ins Leben gerufen. Ziel der Organisation ist es, freiwillige Spezifikationen für eine offene Architektur zu erarbeiten, die ein hohes Maß an technischer Performance, Verlässlichkeit und Qualität sicherstellt.<sup>17</sup>

Das daraus hervorgegangene Dokument „Digital Cinema System Specification“, der eigentliche DCP-Standard, geht genauestens auf Bild-Kodierung und Spezifikationen zum DCP ein, macht jedoch keine detaillierten Angaben zum Ton. Es wird lediglich darauf verwiesen, dass dem DCP 16 Audiokanäle zur Verfügung stehen sollen, dass das unkomprimierte Datei-Format Broadcast Wave verwendet werden soll und dass im Kino mindestens eine 5.1-Anlage oder optional eine 7.1-Anlage installiert sein muss. Die genaue Beschriftung und Zuordnung der Kanäle seien im SMPTE Standard "428-3: D-Cinema Distribution Master Audio Channel Mapping and Channel Labeling" definiert. Im Kapitel „Audio Characteristics“ wird darauf verwiesen, dass der SMPTE Standard "428-2: D-Cinema Distribution Master - Audio Characteristics" einzuhalten sei.<sup>18</sup>

In diesem Dokument wiederum, welches für 50 \$ bei der SMPTE eingesehen werden kann, wird in vier knappen Paragraphen definiert, dass „Audio D-Cinema Distribution Master“ eine Bit-Tiefe von maximal 24 Bit und eine Samplingfrequenz von 48 oder 96 kHz aufweisen sollen, dass 16 Fullrange-Kanäle unterstützt werden sollen und dass digitale Ein- und Ausgänge ein Referenzlevel von  $-20 \text{ dB}_{\text{FS}}$  aufweisen sollen.<sup>19</sup>

Dieser letzte Hinweis deutet also auf die Verbindung zum ursprünglichen Dolby Standard hin. An keiner Stelle wurde von der DCI vermerkt, dass Kinos nach SMPTE RP200 eingemessen werden müssen. Es gibt weder Hinweise, welche Schalldruckpegel bei Referenzlevel von  $-20 \text{ dB}_{\text{FS}}$  erzeugt werden sollen, noch, wie die

Kinoanlage entzerrt wird, noch, welche Dialog- bzw. Maximalpegel eingehalten werden sollen.

Die gegenwärtige Audio-Standardisierung und Regulierung liegt also de facto noch immer bei Dolby, da weiterhin 35mm-Kopien und dementsprechend Printmaster erstellt werden, die anschließend dann auch als Audio D-Cinema Distribution Master weiterverwendet werden. Ab dem Zeitpunkt, wo komplett auf digitale Projektion umgestellt wird, fällt jedoch auch dieser Schritt weg, wodurch nach dem aktuellen Stand ein Kontrollvakuum entsteht.

### ***1.3.3. Werbung und Trailer***

Für Werbespots und Trailer vor dem Langfilm hat sich die Filmwirtschaft auf  $L_{EQ(M)}$ -Obergrenzen geeinigt. Wie dieser Standard und die Regulierung im Einzelnen funktioniert, wird in Kapitel 2.2. erläutert.

## **1.4. Produktion**

### ***1.4.1. Film***

Als Element des Kinosystems besteht die Produktion ihrerseits wieder aus mehreren Gewerken. Die zwei Hauptinteressen, einen subjektiv richtigen und dramaturgisch sinnvollen Lautstärkeverlauf zu wählen sowie ein Produkt mit einer Lautstärke zu erzeugen, das auf dem Markt erfolgreich ist und mit vergleichbaren Filmen konkurrieren kann, sind zunächst keinem dieser Gewerke eindeutig zuzuordnen, jedoch unter Regisseuren, Produzenten, Verleihern, Tongestaltern, Komponisten, Editoren und anderen verschieden gewichtet. Im Folgenden soll für Alle zunächst vereinfachend - entgegen der sonst gebräuchlichen Bedeutung - die Bezeichnung „Filmemacher“ stehen, auch wenn einige Aspekte eher jenen Gewerken zuzuordnen sind, die aktiv an der Mischung beteiligt sind und andere wiederum eher jenen, die den Film verkaufen.

Wo liegt also genau der Unterschied zwischen den Interessen der Filmemacher und denen der Konsumenten? Warum beschweren sich die einen vermehrt über zu leise und die anderen vermehrt über zu laute Pegel? Wenn die Filmemacher eine subjektiv „richtige“ Lautstärke anstreben, dann muss diese doch zwangsläufig auf ihrer eigenen Erfahrung als Konsument beruhen, dann muss sich doch die Interessenkette an dieser Stelle schließen. Die in der Einleitung beschriebene „unüberbrückbare Differenz“ steht im Widerspruch dazu, dass die Filmschaffenden selbst Teil des Publikums sind. Wie verändert sich also die Wahrnehmung der Filmschaffenden im

Bezug auf Kinolautstärke, wenn sie aus ihrem Dasein als Kinogänger heraustreten und selbst Filme machen? Um diese Frage beantworten zu können, muss die Wahrnehmung des Filmtons genauer untersucht werden.

Der französische Filmwissenschaftler und Komponist Michel Chion beschreibt die audiovisuelle Kino-Erfahrung als „Given-to-see“ sowie „Imposed-to-hear“. Der Zuschauer kann seinen Blick demnach auf das ihm gegebene, durch die Leinwand begrenzte Bild richten. Die Ton-Wahrnehmung ist hingegen ein „Ausgesetzt-Sein“, das dem Zuschauer allein durch die Physiologie der Ohren, welche nicht wie Augen geschlossen werden können, durch die Omnidirektionalität des Filmtons, aber auch durch das Fehlen eines auralen Trainings in unserer Kultur ein Selektieren bzw. Ausfiltern von Elementen erschwert. Während sich also ein bewusstes Hören beim Menschen wenig weit entwickelt hat, ist ein unbewusstes Hören unausweichlich. In diesen „Kurzschlüssen zur Wahrnehmung“ sieht Chion das wesentliche Potential des Hörens zur semantischen Manipulation bei der audiovisuellen Wahrnehmung.<sup>20</sup>

Hinzu kommt ein Phänomen, das Chion „Added Value“ (Mehrwert) nennt: Geräusche in einer Szene werden vom Zuschauer wahrgenommen und anschließend als Attribute dem Bild zugeordnet, was bei ihm den Eindruck erzeugt, er sehe etwas, das er in Wirklichkeit „audio-sieht“. So glaubt der Betrachter in einer von wenigen Komparsen bespielten Straßenkreuzung im Film „Blade Runner“ eine tosende und übervölkerte Metropole zu sehen, welche er in Wahrheit vor allem hört.<sup>21</sup>

Wenn also der Filmtone nicht unmittelbar in Erscheinung treten soll, sondern seine dramaturgischen Funktionen unterbewusst erfüllt und beim Zuschauer gar nicht erst den Bewusstseinsgrad erreicht, in dem er als eigenständiges Ereignis und nicht als Teil der Geschichte erkannt wird, dann kann im Umkehrschluss vom Publikum auch nicht erwartet werden, dass es sich seiner Teil-Abwesenheit bewusst wird. Eine Teil-Abwesenheit des Filmtons besteht z.B. dann, wenn der Film so leise wiedergegeben wird, dass subtile Tonelemente von Störgeräuschen wie Popcorn-Rascheln verdeckt werden.

Ein ungeschriebener Vertrag zwischen Filmschaffenden und Publikum klammert bisweilen den szenischen Dialog von dieser Regel der unbewussten Wahrnehmung aus. Der Dialog hat eine Sonderstellung unter den Tonelementen und bietet Zuschauern oftmals eine gewisse verlässliche Konstante. Das Gespräch zweier am Horizont spazierender Figuren beispielsweise, wird dem Publikum meistens genauso wenig vorenthalten wie eines, das unter natürlichen Bedingungen gar nicht

verständlich wäre, etwa die Stimme am anderen Ende der Telefonleitung. Sollte das nicht der Fall sein, dann ist das eine bewusste künstlerische Entscheidung, die sich dem Zuschauer im besten Falle vermittelt. Gebrochen wird der Vertrag, wenn Dialoge unverständlich sind, z.B. zu leise. Der Zuschauer kann dann sehr wohl unterscheiden, ob damit etwas erzählt wird oder ob es sich um einen „technischen“ Fehler handelt. Er greift dann, wenn er die Möglichkeit hat, zur Fernbedienung. Eine subtile Ton-Atmosphäre im Film, die bei niedriger Wiedergabelautstärke nicht mehr zur Geltung kommt, kann diese Reaktion bei ihm nicht hervorrufen. Und nochmals: Das kann einem Zuschauer auch nicht vorgeworfen werden, denn es ist gewissermaßen die Aufgabe solcher subtiler Tongestaltungs-Elemente, sich der Aufmerksamkeit des Publikums zu entziehen.

Hier liegt dann auch der entscheidende Unterschied zum Filmemacher, denn dieser kennt seine Mittel. Er kennt jedes einzelne Element und sieht sein Werk erst dann als vollständig, wenn sich alle von ihm eingesetzten Gestaltungsmittel auch übertragen, auch und vor allem diejenigen, die sekundär wirken und so als „Added Value“ die Wahrnehmung von Bild und Geschichte verändern. Wie bereits angedeutet, sind einige von diesen Elementen viel empfindlicher gegen leisere Wiedergaben, da sie nicht wie Sprache im Bereich eines Referenzlevels gemischt werden, sondern wesentlich tiefer, gelegentlich nahe der unteren Grenze der Kinodynamik. Definiert man also eine Komfort-Zone für Filmemacher, dann fällt diese zwangsläufig symmetrischer aus: Sie verhält sich nach oben ähnlich der von Konsumenten, begrenzt durch die obere Akzeptanz- und Schmerzgrenze. Nach unten muss sie weniger tolerant sein, da eine Wiedergabe nicht erst dann „ärgerlich leise“ ist, wenn die Sprache nicht mehr verständlich bzw. an Alltags-Erfahrungen gemessen zu leise ist, sondern bereits dann, wenn durch die kleinere Dynamik die ersten leise gemischten Teile der Mischung in der Verdeckung verschwinden oder sich nicht mehr genügend vermitteln.

Fraglich bleibt noch, ob sich die Ideallautstärke, absolut betrachtet, zwischen Filmemachern und Zuschauern unterscheidet. Ich denke, dass diese Frage grundsätzlich verneint werden muss, auch wenn die Realität durch psychische Faktoren verzerrt wird.

Filmemacher bewerten Kinolautstärke anders, wenn sie Screenings ihrer eigenen Arbeit besuchen. Es geht dann nicht mehr primär um eine persönliche Ideallautstärke; vielmehr sind sie gedanklich oft bei den übrigen Zuschauern, die ihre Arbeit bewerten. Abhängig vom Stress, der dadurch in ihnen ausgelöst wird,

wollen sie besonders deutlich zeigen, wie großartig der Film ist oder was sie an Arbeit hineingesteckt haben (z.B. Tongestalter) und bemühen sich um eine laute Wiedergabe. Der US-amerikanische Mischtonmeister Tom Hambleton (Chicago Overcoat, Dawning) beschreibt das folgendermaßen:

*„(...) psychology comes into play. People get anxious and want to FEEL the mix, so they turn it up, sometimes too much. When they feel more relaxed, the regular playback seems like enough, or even too much. I notice this even myself.“<sup>22</sup>*

Entscheidend ist also, wie sicher die Filmemacher ihrer eigenen Arbeit gegenüber stehen. Ein Verlangen nach lauterer Wiedergabe kann somit als Kompensation mittelmäßiger Vorarbeit gesehen werden. So versuchen beispielsweise Berufsanfänger im Ton Department oft gleich zu Anfang einen Soundtrack zu erstellen, welcher so eindrucklich klingt wie der eines großen Blockbusters und mischen dementsprechend laut ohne Rücksicht darauf zu nehmen, ob die Elemente wirklich stimmig sind und ob eine laute Passage überhaupt dramaturgisch sinnvoll ist; die Qualität wird zu erzwingen versucht. Mittelmäßig kann aber auch der Film an sich sein. Ein besonders lautes Screening bei der Premiere kann dann einer von vielen Versuchen sein, darüber hinwegzutäuschen, den Zweifel gewissermaßen zu übertönen. Die Wiedergabe ist die allerletzte Möglichkeit der Filmemacher, Einfluss auf die Wirkung ihres Werks zu nehmen.

Das Kalkül „lauter ist besser“ hat sich aber auch als marktwirtschaftliche Tendenz etabliert und kann nicht einfach als Mythos abgetan werden. Psychologisch wird der bloße Unterschied zwischen leise und laut durch menschliche Erfahrungswerte mit klein und groß, schwach und stark, harmlos und gefährlich, mild und intensiv, aber auch fern und nah verknüpft (im Übrigen ist der bloße Schalldruckpegel nur einer von vielen Dimensionen, um ein Element tatsächlich größer bzw. intensiver oder näher wirken zu lassen). Was für den direkten Vergleich von unmittelbar auf einander folgenden Inhalten zutrifft, kann jedoch nicht einfach auf absolute Lautstärken übertragen werden. Während ein Ereignis, das von leise zu laut wechselt, sich augenblicklich intensiviert, darf nicht angenommen werden, dass ein anhaltend lautes Ereignis auch dauerhaft intensiv bleibt. Das Gehör gewöhnt sich bei Langfilmen an die gegebenen Hörbedingungen und setzt Geräusche dann in Bezug dazu. Ein Film, der mit der Dynamik zwischen leise und laut arbeitet, muss und wird diesen dramaturgischen Verlauf also auch bei einer abweichenden Wiedergabelautstärke noch transportieren.

Am Ende steht beim Filmemacher noch das berechtigte Interesse, dass die Zuschauer sein Werk so wahrnehmen, wie er es geschaffen hat. Nicht selten wird deshalb seitens der Tongestalter die radikale Forderung gestellt, den Dolby Regler am Kinoprozessor ganz abzuschaffen und somit eine Kinoauswertung auf Referenzlautstärke zu erzwingen.<sup>23</sup>

#### **1.4.2. Werbung und Trailer**

Was für die Wahl der Lautstärke von Langfilmen gilt, kann nicht ohne weiteres auf die Produktion von Werbungen und Trailern übertragen werden, da diese viel kürzere Programmdauern haben. Ein Spot, der leiser, dafür aber dynamischer gemischt wurde als andere, kann sich nur schwer durchsetzen, da er in einem Block mit vielen weiteren gezeigt wird; ein Phänomen das auch bei Kurzfilm-Blöcken auf Filmfestivals auftritt. Die Lautstärke kann bei jedem Programmwechsel springen und dann augenblicklich Eindrücke wie „stärker“ oder „schwächer“ beim Publikum hervorrufen (Kapitel 1.4.1.). Während ein Zuschauer sich also innerhalb von Langfilmen an die absolute Lautstärke gewöhnt und laute sowie leise Passagen als dramaturgischen Verlauf erfährt, stuft er bei verschiedenen lauten Trailern im unmittelbaren Vergleich die beworbenen Filme möglicherweise als entsprechend besser oder schlechter ein. Die von der TASA festgelegten  $L_{EQ(M)}$ -Obergrenzen (Kapitel 2.2.) werden seitens der Produktion daher in aller Regel vollständig ausgereizt. Damit steht bei Werbungen und Trailern endgültig der Konkurrenz-Aspekt im Vordergrund. Während bei Langfilmen noch darauf spekuliert werden kann, dass sie auf Referenzlevel wiedergegeben oder zumindest individuell geregelt werden, ist spätestens bei der Mischung von Spots entscheidend, wie laut die Konkurrenz mischt.

Interesse der Auftraggeber, also werbenden Firmen und Verleihern, ist entgegen der üblichen Handhabung natürlich eine Lautstärke, die von den Vorführern nicht manipuliert wird, da hierdurch der Werbe-Effekt gemindert wird.

### **1.5. Zwischenbilanz**

Wenn im Bestreben um ein ideales System zunächst die Interessen von Produktion und Konsum verglichen werden, sollten unnütze Faktoren davon aussortiert werden. Auf Seiten des Konsums ist also eine Lautstärke sinnvoll, welche dem Durchschnitt der subjektiven Ideallautstärke der Kinozuschauer entspricht und nicht einer Mittelung zwischen den Toleranzgrenzen, da letztere die jeweils

empfindlichsten Zuschauer bevorzugt (Kapitel 1.1.2.). Auf Seiten der Produktion müssen die psychologischen Effekte im Bezug auf die eigene Arbeit ausgeklammert und somit unrealistische Erwartungen an die Wiedergabe eigener Werke aufgegeben werden. Die Vorstellung, durch bloßes „Pushen“ der Lautstärke - und wenn die Grenzen der Aussteuerung bereits erreicht sind, durch das Steigern der Lautheit - sich einen Vorteil auf dem Markt zu verschaffen, muss entkräftet werden.

Das Dolby System mit dem daraus hervorgegangenen SMPTE Standard RP200 kann bis heute keine zufriedenstellende Lautstärke-Kontinuität schaffen. So genial die Idee auch ist, rein nach Gehör mischen zu lassen und damit die gestalterische Freiheit bei den Filmmachern zu belassen, so sehr ist sie auch dazu verdammt, am menschlichen Fehler zu scheitern. Dieser menschliche Fehler - im Sinne von beliebiger Abweichung - liegt sowohl bei den Filmmachern, als auch bei den Filmvorführern, wie auch beim Publikum. Er bezeichnet ganz einfach die subjektive und damit unterschiedliche Auffassung von dem, was richtig ist. In der Uneindeutigkeit liegt dann auch der Kern des gegenwärtigen Problems.

Der Filmauswertung und somit den Kinobetreibern und Filmvorführern kann die letztendliche Kontrolle der Lautstärke nicht entzogen werden, sie würden Kinoprozessoren ohne Regler ohnehin nicht kaufen. Ein verbessertes System sollte ihnen stattdessen wieder eine Verlässlichkeit bieten. Ein Kinovorführer muss wissen, was er zu erwarten hat, wenn er ein bestimmtes Level am Kinoprozessor einstellt, und nur wenn sich diese Erwartung auch verlässlich bestätigt, werden Vorsichtsmaßnahmen unnötig. Möglich ist das nur durch eine technische Norm, einen Standard, der sicherstellt, dass sämtliche Kinoprogramme in ihrer subjektiv empfundenen Lautstärke vergleichbar sind. Nur dann kann ein Kino sich bewusst dazu entscheiden, seine Filme auf einem von ihm definierten Level zu projizieren. Ein Kino kann sich dazu entscheiden, aufgrund schlechter Schalldämmung oder um eine bestimmte, z.B. ältere Zielgruppe anzusprechen, eine kategorisch gemäßigte Wiedergabe zu wählen oder es kann auf einer definierten Referenzlautstärke projizieren und das auch als Qualitätsmerkmal vermarkten. Möglicherweise können nach britischem Vorbild kindgerechte Vormittags-Vorführungen etabliert werden, sogenannte „Parents and Baby Screenings“,<sup>24</sup> bei denen die Lautstärke um ein definiertes Maß vom Vorführer gemindert wird.


























Eine solche Spartenbildung könnte auch beim Konsumenten eine neue Verlässlichkeit schaffen. Wird sich ein Kinozuschauer seiner persönlichen Lautstärke-Präferenz bewusst und beginnt er sich an lauten oder leisen



Vorführungen derart zu stören, dass sie gar zur Vermeidung von Kinobesuchen führen, dann kann die einzige Lösung nur darin bestehen, dass es bestimmte Kinos gibt, von denen er weiß, dass sie kontinuierlich und somit verlässlich auf einer für diesen individuellen Zuschauer passenden Lautstärke wiedergeben, während bestimmte andere Kinos andere Zielgruppen ansprechen. Ein solches Spartenmodell kann sich dann wirtschaftlich bewähren, wenn es nicht nur das gesamte Spektrum der Nachfrage berücksichtigt, sondern sich auch durch diese dimensioniert. Beispielsweise kann das „Cinestar Original“ in Berlin mit fast zehn Millionen Touristen jährlich in der Hauptstadt rechnen<sup>25</sup> und setzt deshalb auf eine Sparte, ein Angebot von ausschließlich Originalfassungen. Die Nachfrage ist ausreichend hoch und durch abnehmendes Interesse an Synchronfassungen tendenziell steigend. Die Spezialisierung ist auch deshalb so erfolgreich, da sie sich dem Konsumenten vermittelt: Der Hinweis „Original“ steckt bereits im Namen des Kinos. Ein Kinogänger weiß, was er zu erwarten hat und entscheidet dann gezielt nach seinen persönlichen Präferenzen. Eine Transparenz ähnlich den Angaben wie „OV“ oder „O.m.U.“<sup>26</sup> auf Eintrittskarten wäre auch für Wiedergabelautstärken zweckdienlich. Und wenngleich solche expliziten Level-Angaben seitens der Kinos unrealistisch erscheinen, so kann Transparenz doch auch bedeuten, dass die Marketing-Strategie des Kinos eine Wiedergabe für „gemäßigtes Publikum“ oder eben „Sound Enthusiasten“ suggeriert ohne diese explizit so zu benennen. IMAX Kinos z.B. bieten superlative Bild- und Tonbedingungen und vermarkten diese auch durch entsprechende Trailer und Hinweise zu Beginn der Vorführung.

Damit würde also endgültig von der kategorisch gleichen Abhörlautstärke in Kinos abgerückt, die sich ohnehin nicht durchsetzen konnte. Paradoxerweise bestünde gerade darin aber wieder eine Chance, dass auch Filmemacher besser einschätzen könnten, in welchem Lautstärke-Bereich ein Film sein Publikum erreichen wird. Anstatt zu versuchen, ein sich bewegendes Ziel zu treffen<sup>27</sup> und durch Überkompensation Dynamik zu verlieren, hätten sie einen definierten und sinnvollen Rahmen, in dem sie arbeiten könnten. Anders also als auf einer definierten Abhörlautstärke zu mischen, die jener der Auswertung nicht entspricht, könnte nach einem technischen Standard gemischt werden, gleichwohl wissend, dass diese Mischung auf Premieren und repräsentativen Screenings auch tatsächlich wiedergegeben würde wie während der Mischung, aber nicht künstlich gepusht werden kann und auch dann noch funktionieren muss, wenn sie in kommerziellen Vorführungen aus einer der oben genannten Gründe z.B. definiert leiser

wiedergegeben wird. Wenn ein Standard eine solche Verlässlichkeit gewährleistet, dann könnte sich auch beim Mischtonmeister wieder eine eindeutige Zuordnung der verschiedenen Dolby Level zum subjektiven Eindruck einstellen. Er wüsste, wie ein Film klingt, der nach der Norm gemischt ist, wenn er auf Level 7 wiedergegeben wird und was auf Level 5 zu erwarten wäre, sowohl bei seinen eigenen Arbeiten, aber eben auch bei denen seiner Kollegen. Wenn z.B. die Erfahrung zeigt, dass Filme, die sich explizit an Kinder richten, von Kinos auf niedrigeren Levels wie 5,5 projiziert werden, dann würde ein Mischtonmeister das in der Mischung bereits berücksichtigen, nicht etwa durch lauterer „Gegenmischen“, wie es gegenwärtig häufig der Fall wäre, aber ggf. durch die Wahl einer geringfügig kleineren Dynamik. Um weiterhin zu ermöglichen, dass Kinogänger einen Film genauso hören wie die Filmemacher während der Mischung, muss die Sonderstellung der Referenzlautstärke erhalten bleiben. Und wenn dabei nicht mehr die Gefahr besteht, dass überkompensierte Mischungen viel zu laute Wiedergaben im Kino erzeugen, kann der Dolby Regler auch bewusst auf Level 7 belassen werden. Eine Idee wie von Frank Kruse (Sounddesigner zu Cloud Atlas, The International, Das Parfum) „besonders gut gewartete Kinos, die auf 7 abspielen“ auszuzeichnen,<sup>28</sup> würde dem Konzept von Kinobedingungen gleich Mischungsbedingungen wieder Rechnung tragen und zugleich die thematisierte Transparenz beim Kinoangebot fördern. Um die Interessen der Kinobetreiber und Vorführer weiter zu berücksichtigen, müssten Werbe- und Trailermischungen demnach der gleichen Norm oder einer bedingt durch kürzere Dauern und kategorisch dichtere Dramaturgie ggf. angepassten Version dieser Norm entsprechen, die sicherstellt, dass die wahrgenommene Lautstärke der des Hauptfilms entspricht. Da die analoge Projektion zunehmend abgeschafft wird und somit auch der Einfluss von Dolby schwindet, müsste zur Umsetzung weiterhin eine Standardisierungs- sowie Regulierungsinstanz etabliert werden. Wie ein solcher Standard aussehen und wie er technisch umgesetzt werden könnte, soll im Folgenden untersucht werden.

- 
- <sup>1</sup> im Folgenden vgl.  Riedmiller, Lyman, Robinson: (Dolby Laboratories): Intelligent Program Loudness Measurement and Control - What Satisfies Listeners?, S. 5f
- <sup>2</sup> vgl.  Benjamin, Eric: Comparison of Objective Measures of Loudness Using Audio Program Material
- <sup>3</sup> vgl.  Pearsons, Bennet, Fidell: Speech Levels in Various Noise Environments
- <sup>4</sup> Am Kinoprozessor entsprechen 0,5 Schritte von 0-4 je 10dB und von 4-10 je 3/5dB
- <sup>5</sup>  $(|10,8| + |-19,1|) / 2 - 19,1 = -4,15$  [dB]
- <sup>6</sup> vgl.  EBU: Empfehlung R 128, S. 4
- <sup>7</sup> siehe  Burg, Lothaire (ARTE): Memo für Produzenten & Dienstleister - Lautheit von Fernsehprogrammen
- <sup>8</sup> vgl.  Thom, Randy (Sound-Article-List): The future of hollywood sound... (09.03.09, 01:56)  
„At least one of the largest movie theater chains has an official policy of playing all films approximately six db below the calibrated level.“
- <sup>9</sup> vgl.  Dolby Laboratories: Technical Guidelines for Dolby Stereo Theaters, S. 12ff
- <sup>10</sup> vgl.  Dolby Laboratories: Dolby Atmos Cinema Technical Guidelines, S. 19
- <sup>11</sup>  Dolby Laboratories: CP750 Digital Cinema Processor Manual, S. 3
- <sup>12</sup> vgl.  DTS: Installation and Operation Manual DTS-6D, S. 9f
- <sup>13</sup> vgl.  Sony: SDDS Print Master Guidelines, S. 5
- <sup>14</sup>  Zernicke, Jerry (THX, Director of Professional Services): Cinema Playback Levels
- <sup>15</sup>  Lerch, Christian (Dolby Laboratories, Senior Sound Consultant, Content Services): Lautheitsmessungen bei Kinofilmen
- <sup>16</sup> vgl.  SMPTE: Recommended Practice RP 200:2012, S. 11
- <sup>17</sup> vgl.  DCI: About DCI
- <sup>18</sup> vgl.  DCI: Digital Cinema System Specification, S. 26f
- <sup>19</sup> vgl.  SMPTE: SMPTE Standard 428-2-2006 D-Cinema Distribution Master - Audio Characteristics, S. 2
- <sup>20</sup> vgl.  Chion, Michel: Audio-Vision, S. 33f
- <sup>21</sup> vgl.  Chion, Michel: Film, A Sound Art, S. 212f
- <sup>22</sup>  DUC: DCP cinema volume discrepancy (minister, 22.03.2012, 11:33)  
„Auch Psychologie spielt eine Rolle. Leute werden nervös und wollen die Mischung FÜHLEN, also drehen sie auf, manchmal zu viel. Wenn sie sich entspannter fühlen, erscheint die normale Wiedergabe genügend oder sogar zu laut. Ich merke das auch bei mir selbst.“
- <sup>23</sup> vgl.  DUC: Dolby 7. Is everybody mixing at this level? (airon, 14.05.11, 16:14): „One has to wonder if the Dolby unit should have a damn knob at all.“
- <sup>24</sup>  Londonnet: Movies with parents and baby screenings currently showing in the UK
- <sup>25</sup> vgl.  Berliner Morgenpost (15.02.2012)
- <sup>26</sup> Original Version sowie Original mit Untertiteln
- <sup>27</sup> vgl.  DUC: Dolby 7. Is everybody mixing at this level? (JKD99, 13.05.11, 08:52)  
„IMHO, we need to keep the standard that 7 on the cinema processor =85db; otherwise you're chasing a moving target.“
- <sup>28</sup> vgl.  BVFT: Abhörlautstärke bei Trailermischungen (Frank Kruse, 25.02.09, 20:54)

## 2. Lautheits-Messungen - Verfahren und Anwendungen

### 2.1. Objektive Lautstärke-Messmethoden

Traditionell kann die Lautstärke als Lautstärkepegel in phon angegeben werden (logarithmisch) oder als Lautheit in sone (linear). Da es sich dabei aber um psychoakustische Maße handelt, können sie nicht einfach aus dem Schalldruckpegel hergeleitet werden. Ihnen kann sich lediglich angenähert werden und so wurden in der Vergangenheit eine Vielzahl von Lautheits-Messgeräten entwickelt, die sich jeweils unterschiedlicher, sogenannter „objektiver Messmethoden“ bedienen. Als Grundlage dienten dafür nicht nur Hörversuche mit Testpersonen, sondern auch bestehende Erkenntnisse aus der Psychoakustik. Zum einen sind das die spektrale Verdeckung, bei der benachbarte Frequenzgruppen sich pegelabhängig maskieren sowie die zeitliche Verdeckung, durch die Klänge unmittelbar bevor und nach lauterem Schallereignissen maskiert werden. Hinzu kommt die menschliche Frequenzwahrnehmung, die in den Kurven gleicher Lautstärkepegel beschrieben wird, das Impulsverhalten des Ohres sowie die räumliche Wahrnehmung und Einstrahlrichtungs-abhängige Entzerrung des Außenohres.<sup>1</sup>

Zwischen zwei Prinzipien wird unterschieden:

- *Single-Band Algorithmen*

Das Signal durchläuft bei Single-Band Algorithmen lediglich einen Weg. Um dem Zeitverhalten des Hörens Rechnung zu tragen, nämlich dass kurze Spitzen oder kurzweilige laute Passagen das Lautstärke-Empfinden weniger beeinflussen, wird häufig eine RMS (Root Mean Square) Messung über einen längeren Zeitraum, z.B. über die Dauer eines Programmes, durchgeführt, was als  $L_{EQ}$ -Messung (Langzeit äquivalenter durchschnittlicher Lautstärkepegel) bezeichnet wird. Die menschliche Frequenzwahrnehmung wird dabei gewöhnlich durch eine Gewichtung simuliert, die sich typischerweise aus den Kurven gleicher Lautstärkepegel ableitet. Das bedeutet auch, dass solche Gewichtungen nur für einen bestimmten Schalldruckpegel-Bereich ausgerichtet sind. Besonders häufig kommen Messungen mit A-Gewichtung - ursprünglich nur für niedrige Pegel - und C-Gewichtung für höhere Pegel mit den jeweiligen Einheiten  $dB_A$  und  $dB_C$  vor (siehe Abbildung 3).<sup>2</sup> Weitere Gewichtungen sind die B-Kurve für mittlere und die D-Kurve für besonders hohe Schalldruckpegel wie Fluglärm, sowie die

RLB-Kurve, die eine um den Tiefenbereich erweiterte B-Kurve darstellt und die M-Kurve zur von Dolby entwickelten Bewertung von Filmen (Kapitel 2.2.2.).

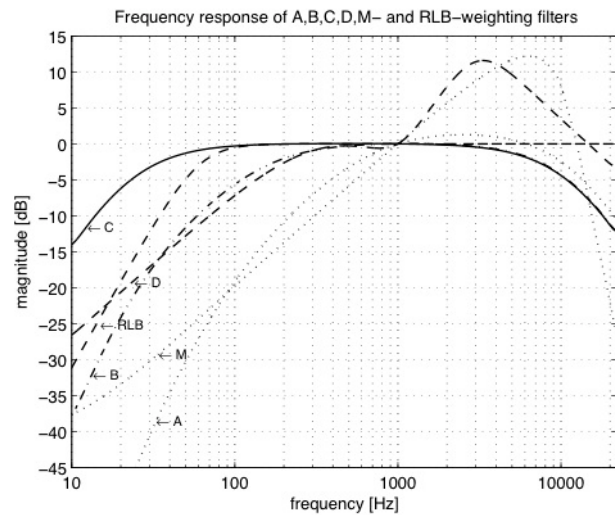


Abbildung 3: Gewichtungs-Kurven

- **Multi-Band Algorithmen**

Bei Multi-Band Algorithmen wird ein Signal nach einer Vorentzerrung in mehrere Bänder gefiltert, die ihrerseits häufig den „kritischen Bandbreiten“ nach Fletcher<sup>3</sup> entsprechen oder diesen mehr oder weniger ähneln, z.B. durch drittel, halbe oder ganze Oktaven. Die Bänder werden separat ausgewertet und anschließend zu einem Lautheitswert zusammengeführt. Dadurch werden, anders als bei Single-Band Algorithmen, auch Maskierungseffekte zwischen Frequenzen berücksichtigt. In den Standards ISO 532 A und B wurde diese Methode von Stevens und Zwicker beschrieben und anschließend immer wieder überarbeitet und verfeinert.<sup>4</sup> Wenngleich zahlreiche Echtzeit-Anwendungen von Multi-Band Algorithmen existieren, fand diese Methode möglicherweise aufgrund ihrer hohen Rechenanforderungen nahezu keine Anwendung in gesetzlichen Immissionsschutz-Bestimmungen oder Aussteuerungs-Normen.<sup>5</sup>

Auf der Suche nach dem optimalen Lautstärke-Messgerät muss neben dem zu erwartenden Wiedergabepegel (Kurven gleicher Lautstärkepegel) vor allem auch nach dem zu messenden Audio-Material gefragt werden. Während frühe Forschungen vor allem Testtöne und Bandpassrauschen als Stimuli benutzten um Erkenntnisse zur Lautstärke zu sammeln, wurde in den letzten Jahrzehnten verstärkt Forschung zu Messungen bei gemischten Testsignalen betrieben, wie sie z.B. bei Fernsehprogrammen auftauchen. Die Motivation für diese Versuche ähnelte nicht selten derer der vorliegenden Arbeit: Ziel war die Konzeption von Algorithmen für die Auswertung von Klanggemischen aus Sprache, Musik und Geräuschen für eine Lautstärke-Anpassung in verschiedenen medialen Bereichen, allen voran TV. Drei solcher Bemühungen weisen Parallelen zum Bestreben dieser Arbeit auf und haben

sich als Standard etabliert. Sie sollen im Folgenden untersucht werden um festzustellen, welche Messmethoden und gewonnenen Erkenntnisse für eine Kinoaussteuerungs-Norm in Frage kämen.

## 2.2. TASA Standard nach $L_{EQ(M)}$

### 2.2.1. Initiative

Im März 1997 wurde auf der SMPTE Film Konferenz ein historisches Papier mit dem Titel „Are Movies Too Loud?“ veröffentlicht, welches in großen Teilen die noch immer aktuelle Problem-Situation zur Kinolautstärke beschreibt.<sup>6</sup> Angeregt durch die National Association of Theater Owners (NATO) wurde die Arbeit von Dolby Laboratories in Auftrag gegeben und von Vice President Joan Allen persönlich verfasst. Erstaunlich kritisch geht er darin mit Filmemachern ins Gericht, die Filme produzieren, die kategorisch und durchgängig laut sind und Beschwerden beim Publikum forcieren:

*„The ‚ride‘ movie is now commonplace from Hollywood, and while the film can be exciting, the phrase ‚cut-to-the-chase‘ no longer has any meaning - the chase frequently starts at the beginning of reel 1!“<sup>7</sup>*

Allen schreckt aber auch nicht davor zurück, die Gründe für lauter werdende Filme bei Dolby selbst und den in den neunziger Jahren eingeführten Digitalformaten Dolby Digital, aber auch SDDS und DTS zu suchen. Der gewonnene extra Headroom von nunmehr 20 dB sei ein Mehrwert für die künstlerische Gestaltung der Filme, mache aber auch derartigen Missbrauch überhaupt erst möglich. Er stellt unter anderem fest, dass Mischtonmeister regelmäßig mit Gehörschutz mischen und bezeichnet die gegenwärtige Situation in den späten neunziger Jahren als eindeutig unbefriedigend. Folglich untersucht er die Möglichkeiten, die Lautheit von Mischungen eindeutig beziffern zu können und sieht im  $L_{EQ(M)}$ -Verfahren den richtigen Weg dafür. Die Art von Allens Argumentation macht auch unmissverständlich klar, dass sich Dolby als Entwickler und nicht - wie gelegentlich angenommen - als Regulator versteht. Er sieht in  $L_{EQ(M)}$ -Messungen ein geeignetes Instrumentarium, halte es aber für anmaßend, selbst Maximalwerte festzulegen; dies sei Aufgabe der „künstlerischen Community“. Als „erster Schritt zur Vernunft“ erscheine es wünschenswert, Druck auf Produzenten von Werbungen und Trailern auszuüben,  $L_{EQ(M)}$ -Werte zu senken. Das klingt, als sei auch eine Regulierung der Filme selbst denkbar, auch wenn der „Hauptangriff“ auf das Referenzlevel im Kino laut Allen von Werbungen und Trailern ausgehe. Er stellte in seinen Untersuchungen

1997  $L_{EQ(M)}$ -Werte bei Trailern fest, die zum Teil 20 dB über denen der Filme lagen, die sie bewarben.

### **2.2.2. Verfahren**

Allen bediente sich für die Konzeption eines Film-Lautstärke-Messgerätes zweier bereits bekannter Komponenten: Der RMS-Integration über die Zeitspanne eines Programms ( $L_{EQ}$ ) und der CCIR-468-Gewichtungskurve, die ursprünglich zur Bewertung von Geräuschpegeln bei Aufnahmesystemen standardisiert wurde und ihm durch ihre Anhebung bei 2-6 kHz günstiger erschien als die A-Gewichtung. Die Kurve wurde durch ein 5,6 dB Offset auf 2kHz normalisiert. Alle Kanäle werden entsprechend ihrem Tonformat und den dazugehörigen Schalldruckpegeln im Kino (SMPTE RP200) aufsummiert, bevor die Messung durchgeführt wird. Der resultierende Wert repräsentiert den Schalldruckpegel, den der Film bei Wiedergabe über ein entsprechendes System erzeugen würde.<sup>8</sup> Das Verfahren, das aufgrund seines Anwendungsbereichs das Kürzel „M“ (für Movie) trägt, wurde erstaunlicherweise kaum durch Studien verifiziert. Allen begründet die Wahl der CCIR-Kurve lediglich damit, dass diese praktischerweise bereits vorhanden sei und bessere Ergebnisse als die A-Kurve liefere. Es gebe keine technische Parallele zwischen Filmtönen und Geräuschpegeln bei Aufnahmesystemen, also dem ursprünglichen Anwendungsbereich der CCIR-Kurve.<sup>9</sup>

Bei späteren Studien des konkurrierenden Entwicklers TC Electronic und der Universität Aarhus schnitt das  $L_{EQ(M)}$ -Verfahren zwischen zwölf Lautheitsmessgeräten zur Bewertung von Musik und Sprache mit Klasse 4 von 4 schlecht ab.<sup>10</sup>

### **2.2.3. Umsetzung**

Im April 1999, zwei Jahre nach Allens Studien und der Vorstellung des ersten Dolby  $L_{EQ(M)}$ -Meters, veröffentlichte die Trailer Audio Standard Association (TASA), eine Vereinigung aus Filmstudios, Entwicklern (Dolby, DTS, Sony) und weiteren Organisationen (NATO, MPAA), den TASA Standard, der einen maximalen  $L_{EQ(M)}$ -Wert von 87 für Trailermischungen vorschrieb. Dieser Wert wurde später zweimal gesenkt und liegt seit 2001 bei  $L_{EQ(M)}$  85.<sup>11</sup> TASA sah in dem Dokument eine regelmäßige Überprüfung und ggf. Anpassung des Standards vor, sollte sich die Empfehlung in der Praxis nicht bewähren. Der TASA Standard hat sich anschließend international etabliert und wurde von weiteren nationalen Organisationen und Institutionen übernommen. Zur Umsetzung der Vorschrift wurde ein

Zertifikate-System eingeführt. Demnach soll eine sog. „I.A.E.“ (Independent Audio Engineering Company) die Konformität zum Standard kontrollieren und anschließend das TASA Zertifikat ausstellen. In Deutschland überwacht die FSK (Freiwillige Selbstkontrolle der Filmwirtschaft) neben Altersfreigaben auch die Einhaltung des TASA Standards. Als Einrichtung der Spitzenorganisation der Filmwirtschaft e.V. (SPIO) verpflichtet sie ihre Mitglieder, nur von der FSK geprüfte Produkte öffentlich anzubieten.<sup>12</sup>

Die Screen Advertisers World Association (SAWA) brachte 2000 zudem eine Empfehlung für Kinowerbungen von maximal  $L_{EQ(M)}$  82 heraus.

2004 standardisierte die ISO das  $L_{EQ(M)}$ -Verfahren international als Methode zur Messung von wahrgenommener Lautstärke bei Kinofilmen.<sup>13</sup>

#### **2.2.4. Erfolg**

Während sich die Empfehlungen von TASA und SAWA durchsetzten und das Spannungsfeld zwischen Vorprogramm und Hauptfilm in geringem Ausmaß verkleinern konnten (Kapitel 3.5.1.), blieb der gewünschte Effekt einer nunmehr konsequenten Kinowiedergabe auf Referenzlevel aus. Filmvorführer sehen sich weiterhin dazu veranlasst, Werbungen und Trailer leiser zu projizieren. Mischtonmeister Matthias Lempert sieht darin noch immer den Hauptgrund für inkonsistente Wiedergabelevel im Kino und fordert einen deutlich niedrigeren  $L_{EQ(M)}$ -Maximalwert von mindestens 10 dB weniger, also 75 dB<sub>(M)</sub>. Da „jedes Zehntel dB  $L_{EQ}$  ausgereizt“ werde, sei in der Realität und bei der Wiedergabe auf Referenzlevel jeder Trailer ein „Angriff auf die Gesundheit der Zuschauer“ und würde zu Recht vom Vorführer leiser projiziert. An anderer Stelle sagt er „(...) mixing trailers is not mixing but distorting...“<sup>14</sup> und meint damit das Arbeiten an der oberen Aussteuerungsgrenze sowie den übertriebenen Gebrauch von Peaklimitern.<sup>15,16</sup>

Damit kommt ein interessanter Aspekt in die Thematik:

Wenn wie bei Trailern Vorgaben zur maximalen Lautheit gemacht werden, dann geschieht das auch deshalb, weil Vorgaben zur maximalen technischen Aussteuerung (z.B. Maximal-Spannungspegel nach QPPM) umgangen werden könnten und nur zur Folge hätten, dass Mischtonmeister im Bestreben um eine lautere Mischung mehr komprimieren und begrenzen würden. Durch letzteres wird die technische Vollaussteuerung nicht überschritten, aber die Lautstärke gesteigert. Allgemein wird davon ausgegangen, dass der RMS-Wert (wie bei  $L_{EQ}$ -Messungen) ansteigt, sobald bei gleichbleibendem Spitzenpegel Kompression angewandt wird



und dass Vorgaben zur maximalen Lautheit deshalb den Gebrauch von Kompression unnütz erscheinen lassen. Eine Studie aus dem Jahr 2004<sup>17</sup> belegt aber, die Lautstärke könne auch bei gleichbleibendem RMS-Wert durch Kompression gesteigert werden; erst komplexere Lautstärkemesser könnten die Veränderung nachvollziehen.<sup>18</sup> Eine Manipulation mithilfe von Kompressoren und Limitern kann also selbst bei Vorschriften oder Empfehlungen, die auf  $L_{EQ}$ -Messungen beruhen, nicht vollständig ausgeschlossen werden, wenngleich sie dadurch massiv eingeschränkt wird.

Letztlich besteht die Schwierigkeit, einen passenden  $L_{EQ(M)}$ -Maximalwert für Werbungen und Trailer zu finden auch darin, dass die  $L_{EQ(M)}$ -Werte bei Hauptfilmen extrem unterschiedlich und zudem undokumentiert sind.

Es sei hier auch kurz angemerkt, dass Dolby 1997 offenbar erstaunlich gut um die Probleme des gegenwärtigen Kino-Standards Bescheid wusste, aber in den folgenden Jahren keine Initiative ergriffen hat die Lautstärke-Kontinuität zwischen Langfilmen zu verbessern.

## 2.3. AC-3 und Dialog-Normalisierung

### 2.3.1. Verfahren

Im Jahr 1992 führte Dolby Laboratories das Lichtton-Verfahren Dolby Digital ein. Der ihm zugrunde liegende Codec AC-3 (Audio Compression - 3) ist jedoch universeller einsetzbar als nur für 35mm Film. Er kann Mono- bis 5.1-Signale von 16 bis 20 Bit und 32 bis 48 kHz mit Datenraten von 32 bis 640 kBit/s komprimieren. Und er besitzt eine besondere Metadatenstruktur, welche zwar beim Lichtton-Verfahren nicht zum Einsatz kommt, aber für Heimanwendungen eine Möglichkeit zur Lautstärke-Anpassung verschiedener Programme bietet. Genauer gesagt, wird keine Anpassung der Gesamtlautstärke angestrebt, sondern eine sog. Dialog-Normalisierung, bei der Dialoge über verschiedene Programme hinweg mit einer - abgesehen von dramaturgischer Variation - konstanten Lautstärke wiedergegeben werden sollen. Damit ging Dolby auf Erkenntnisse wie jene von Jeffrey C. Riedmiller et al.<sup>19</sup> ein, nach denen Zuschauer ihre Wiedergabe-Lautstärke am Dialog ausrichten (Kapitel 1.1.1.).

Anders als beim  $L_{EQ(M)}$ -Verfahren wird nicht nach einem bestimmten Standard angesteuert. Stattdessen soll im Anschluss an die Mischung eine A-bewertete  $L_{EQ}$ -Messung von durchschnittlichen Dialogpassagen durchgeführt bzw. das Dialoglevel abgeschätzt und der jeweilige Wert als „Dialnorm“ (von -1 bis -31) in den

Metadaten verzeichnet werden. Erst bei der Wiedergabe werden die Metadaten ausgewertet und das Signal durch Absenkung entsprechend dem Dialnorm-Wert normalisiert, sodass am Ende die Dialoge aller Programme durchschnittlich bei  $L_{EQ(A)}$  -31 dB<sub>FS</sub> liegen.<sup>20</sup> Im Allgemeinen wird davon ausgegangen, dass  $L_{EQ(A)}$  zwar für Musik keine befriedigenden Ergebnisse liefert, für Sprache jedoch schon, weswegen es für die Dialog-Normalisierung ausreichend ist.<sup>21</sup>

Zusätzlich zur Dialog-Normalisierung ist eine Wiedergabe-seitige Dynamik-Einengung, die sog. DRC (Dynamic Range Control) vorgesehen. Durch Presets, die ebenfalls beim Mastering des AC-3 Datenstroms als Metadaten gesetzt werden, können passend zum Programm bestimmte Vorgaben gemacht werden: Z.B. werden beim Preset „Film Light“ - sofern die DRC beim Konsumenten eingeschaltet ist - Signalanteile, die 14dB unter dem Dialoglevel liegen, mit einer Ratio von 2:1 angehoben. Ab 1 dB über dem Dialog wird mit 2:1 komprimiert und ab 15dB mit 20:1 limitiert. Werden alle Metadaten richtig gesetzt, ist der AC-3 Standard in der Lage, sehr unterschiedliche Programme aus z.B. Kino oder TV einander anzupassen und noch dazu eine passende Programmdynamik zu gewährleisten.<sup>22</sup>

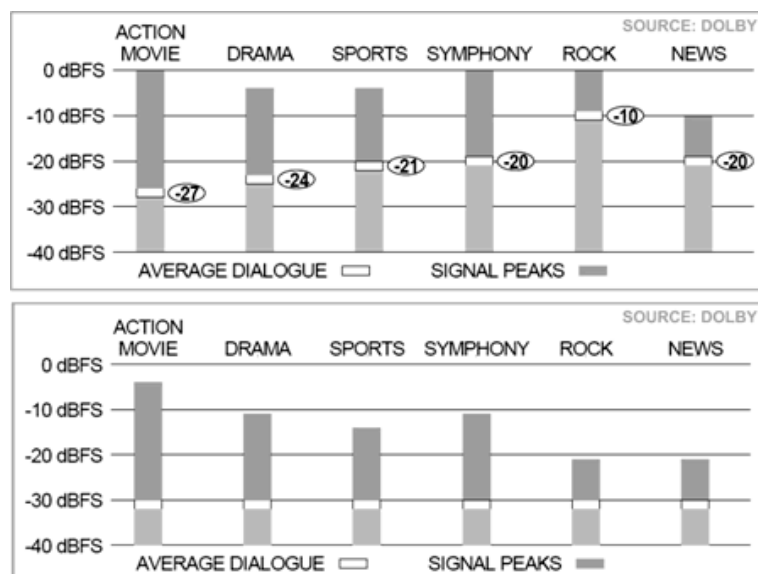


Abbildung 4: Dialog-Normalisierung versch. Programme

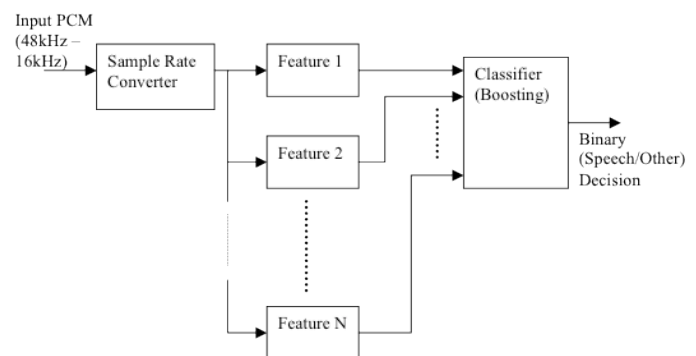
### 2.3.2. Dialogue Intelligence™

Zur Bestimmung des Dialoglevels für den Dialnorm Parameter hatte Dolby 2002 das LM100 Lautheits-Messgerät herausgebracht, das eine A-gewichtete  $L_{EQ}$ -Messung an stereophonen PCM-Signalen oder mehrkanaligen AC-3 bzw. Dolby E Datenströmen durchführt. Das Problem bei dieser Methode ist, dass die Dialog-Passage vom Benutzer ausgesucht wird und so für ein Programm zwangsläufig immer uneindeutige  $L_{EQ(A)}$ -Werte entstehen. Diese Schwachstelle könnte theoretisch auch ausgenutzt werden, um einen möglichst niedrigen Dialnorm-Wert und somit lautere Wiedergaben zu erreichen. Um eindeutige Ergebnisse zu ermöglichen, brachte Dolby deshalb bereits 2003 eine Erweiterung für den LM100 heraus, die Dialogue

Intelligence (im Folgenden DI). Durch diesen Algorithmus soll die Messung immer nur dann eingeschaltet werden, wenn Sprache im Signal vorhanden ist. Dadurch kann über die gesamte Programmdauer gemessen werden, ohne dass Passagen mit Musik oder Soundeffekten in die Bestimmung des Dialnorm-Wertes mit einfließen. Der Algorithmus selbst entscheidet nur, ob gemessen werden soll oder nicht und ist an keine Methode zur Lautheits-Messung gebunden, wengleich im LM100 nach  $L_{EQ(A)}$  gemessen wird.

Die sog. „Speech/Other“ Diskriminierung (Abbildung 5) wandelt zunächst das Eingangssignal in eine Abtastrate von 16 kHz, die für die Erkennung von Sprache als ausreichend erachtet wird.

Anschließend werden vom resultierenden Signal sieben Features (Sprachmerkmale) berechnet, die aufgrund ihrer spezifischen Funktionsweise in verschieden großen Blöcken rechnen. Jeder Block gibt ein Ergebnis von 1 (Sprache) oder 0 (keine Sprache) aus und je nach



**Abbildung 5:**  
**Blockschaltbild Dialogue Intelligence Algorithmus**

Größe bilden eine entsprechende Anzahl Blöcke zusammen ein Frame. Die Framegröße ist für alle Features wieder gleich und beträgt 32768 Samples, was einer Dauer von ca. 2s entspricht (Tabelle 1). Am Ende stehen also für jedes 2s-Frame des Audiosignals die Ergebnisse von sieben Features zur Verfügung, die durch einen abschließenden „Boosting“-Prozess zur endgültigen Entscheidung zwischen „Speech“ oder „Other“ führen.

Folgende Sprachmerkmale werden zur Entscheidung untersucht:

1. Durchschnittliche quadratische  $L_2$ -Norm des gewichteten Spektralflusses:

Bei diesem Prozess wird untersucht, ob das Audiosignal ein sich schnell änderndes Frequenzspektrum aufweist, das für Sprache typisch ist, da sich tonale Elemente abrupt mit geräuschhaften Elementen abwechseln. Bei tonalen, also stimmhaften Fragmenten, finden außerdem häufig innerhalb eines Zeitabschnitts Tonhöhen-Änderungen statt, was sich maßgeblich von Musik und sogar Gesang unterscheidet, bei denen Tonhöhen weniger häufig geändert werden. Diese Berechnung ist mit einer Fehlerrate von 6,67% die sicherste der sieben Features.

- Steigung einer regressiven Linie zur Beschreibung der spektralen Leistungsdichte: Dieses Feature analysiert den Wechsel zwischen positiver und negativer Steigung einer Linie, welche die spektrale Gewichtung zu hohen und negativen Frequenzen nachempfunden. Enthält ein Signal hauptsächlich hohe Frequenzanteile, ist die Steigung positiv, bei hauptsächlich tiefen Frequenzanteilen negativ. Auf Sprache würde hindeuten, wenn die Steigung häufig zwischen positiv (stimmlos, geräuschhaft) und negativ (stimmhaft, tonal) wechselt und häufiger negativ als positiv ist, da stimmhafte Anteile bei Sprache dominieren. Eine symmetrische Verteilung hingegen würde auf Musik oder Geräusche

hindeuten.

- Pausenzählung: Das häufige Auftreten von Pausen, also kurzen Abschnitten von verhältnismäßig geringer Energie kann auf Sprache hindeuten, ist jedoch auch anfällig für impulshafte Geräusche; mit 30,89% hat dieses Feature die größte Fehlerrate.

- Steigungs-Koeffizient der Null-Durchgangs-Rate: Wieder wird hier die Tatsache ausgenutzt,

dass Sprache mehr stimmhafte als stimmlose Anteile enthält, was eine asymmetrische Verteilung hin zu wenigen Nulldurchgängen innerhalb eines Zeitabschnitts nach sich zieht. Bei Musik und Geräuschen ist die Verteilung nahezu symmetrisch.

- Mittel- zu Medianwert Verhältnis der Null-Durchgangs-Rate: Bei diesem Feature wird das gleiche Sprachmerkmal wie unter 4. durch einen Mittel- zu Medianwert beschrieben.

- Kurze Messung der Rhythmik: Hier sollen von den bisher als Sprache erkannten Passagen Elemente wie Rap aber auch Gesang in Popmusik ausgeklammert werden, welche eine eindeutiger Rhythmik aufweisen.

Feature	Block Size (samples)	Block Length (msec)	Blocks per Frame
Average squared $l_2$ -norm of weighted spectral flux	1024	64	32
Skew of regressive line of best fit through estimated spectral power density	512	32	64
Pause count	256	16	128
Skew coefficient of zero crossing rate	256	16	128
Mean-to-median ratio of zero crossing rate	256	16	128
Rhythmic measure	256	16	128
Long rhythmic measure	256	16	128

**Tabelle 1: DI Features und deren Blockgrößen**

7. Lange Messung der Rhythmik: Ein weiteres Ausschlussverfahren von stark rhythmischen Sprachpassagen, wobei dieses hier durch spektrale Gewichtungen und nicht, wie unter 6., durch Mittelwert-freie Sequenz-Werte berechnet werden.

Dolby testete den entwickelten Algorithmus an einer Auswahl von Klangbeispielen aus Fernsehen, Hörfunk und Filmen, wobei durch die Kombination der sieben Features eine Gesamtfehlerrate von 3,36% erreicht wurde. Demnach ist die Wahrscheinlichkeit, dass andere Signalanteile als Sprache identifiziert werden (3,67%), knapp höher, als dass Sprache nicht erkannt wird (2,95%). Die Klassifizierung nimmt lediglich 15-MIPS (15 Mio. Instruktionen pro Sekunde) und 30k Wörter RAM in Anspruch, was eine kostengünstige Hardware-Anwendung und Echtzeit-Berechnung ermöglicht. Der DI-Algorithmus wurde von Dolby bewusst mit einer Latenz konzipiert, die 2 Sekunden nicht überschreitet und steht im übrigen zur lizenzierten Implementierung in Hard- und Software von Drittanbietern als freier Download zur Verfügung.<sup>23,24</sup>

### **2.3.3. Umsetzung**

Nach der Implementierung vom AC-3 Codec ins Dolby Digital Lichtton-Verfahren setzte sich der Standard 1995 in seiner vollendeten Form für Laser Disc und anschließend, nach einem jahrelangen Normenstreit, als das einzig vorgeschriebene Mehrkanal-Format für DVD-Video durch.<sup>25</sup> Damit kamen erstmals die Möglichkeiten zur Dialog-Normalisierung und DRC im Heimbereich zur Anwendung. Die Neuauflagen Dolby Digital Plus und Dolby TrueHD erweiterten dabei die Datenrate bzw. ermöglichten die Datenreduktions-freie Kodierung von 24Bit/96kHz PCM-Audiosignalen z.B. für Blu-ray Disc.

In den USA löste das Advanced Television Systems Committee (ATSC) Mitte 2009 das analoge TV System NTSC durch den ATSC Standard A/53 für Digital-, Kabel-, Satelliten- und terrestrischen Empfang ab. Bereits früh hatte festgestanden, dass dieser Standard den AC-3 Codec (definiert durch ATSC A/52) als einziges Audio Format vorsehen würde. Die neue Fernsehnorm definiert, dass zwingend ein gültiger „Dialnorm“-Wert per Metadaten übertragen wird, der durch eine  $L_{EQ(A)}$ -Messung von Dialogpassagen bestimmt werden soll.<sup>26</sup> Die Federal Communications Commission (FCC) übernahm diese Dialnorm-Pflicht sogar als US-Gesetz für den digitalen Rundfunk. Viele Sender und Produzenten waren sich jedoch unsicher, wie mit dem Dialnorm Parameter zu verfahren sei, weder war man

sich sicher, wie dieser im Einzelnen einzustellen noch wie er eindeutig zu ermitteln sei. Also machten die Sender eigene feste Vorgaben, nach denen gemischt und gemessen, sowie anschließend als AC-3 gemastert werden sollte (z.B. PBS Dialoglevel bei „-24 dB<sub>FS</sub>“, NBC bei „-23 dB<sub>FS</sub>“, gemessen mit L<sub>EQ(A)</sub>). Andere Sender, wie z.B. Discovery Channel, setzten sogar die Messung mit LM100 und DI voraus und machten Vorgaben zu einem „Infinite Dialogue“ Wert zwischen -26 dB<sub>FS</sub> und -28 dB<sub>FS</sub> (Ergebnis einer L<sub>EQ(A)</sub>-Messung mit aktivierter Dialogue Intelligence).<sup>27</sup>

Um die zunehmende Nachfrage nach einer Lösung zur Anpassung von Sendebiträgen zu decken, brachte Dolby den Program Optimizer DP600 heraus, der Lautheits-Werte (darunter z.B. L<sub>EQ(A)</sub> gemessen mit DI) aus einer Vielzahl von Formaten wie AC-3 sowie PCM bestimmt, diese benutzerdefiniert normalisiert und anschließend neu kodiert. Während dieses Verfahren vor allem Sendern die Arbeit erleichtern soll, kam für die Produktion zusätzlich zum LM100 die Software Dolby Media Meter heraus, die ebenfalls mit DI messen kann und, anders als der LM100, 5.1 PCM-Signale unterstützt.

AC-3 ist neben MPEG-2 auch ein mögliches Tonformat für DVB-T terrestrisches Digital-Fernsehen. In Deutschland wurden Mehrkanal-Programme zunächst teilweise mit AC-3 gesendet. Die Praxis wurde aber wegen zu schmaler Bandbreite wieder aufgegeben, und inzwischen halten sich nur noch einige wenige öffentlich-rechtliche Sender die Option offen, mehrkanalig mit AC-3 zu senden.<sup>28</sup> Bei DVB-C (Kabel) und DVB-S (Satellit) sind mehrkanalige Übertragungen mit AC-3 häufiger, allerdings werden Stereo-Programme, also der Großteil aller Sendungen, grundsätzlich mit AAC übertragen, was eine konsequente Dialog-Normalisierung im deutschen TV ohnehin unmöglich macht.

#### **2.3.4. Erfolg**

Wiedergabeseitig lässt sich feststellen, dass die Möglichkeiten für Konsumenten durch Ein- und Ausschalten bzw. Staffeln der DRC eine bevorzugte Wiedergabedynamik zu wählen, eher selten bewusst angewendete Features des AC-3 Standards sind, da sie ein erweitertes technisches Verständnis voraussetzen. Durch Bezeichnungen wie „Tag-“ oder „Nachtmodus“ haben viele Gerätehersteller versucht, den Standard dem Zuschauer näherzubringen. Insgesamt muss die DRC aber als positiv bewertet werden, da sie z.B. Heimkino-Enthusiasten eine größere Dynamik ermöglicht, hingegen eine Default-Wiedergabe ohne benutzerdefinierte

Einstellung meist zu einer komprimierten Dynamik führt, die für durchschnittliche Fernsehzuschauer angemessen erscheint.

Produktionsseitig sind bei der Erstellung von DVD-Videos und Blu-ray Discs Probleme festzustellen. In einer Untersuchung aus dem Jahr 2010 wurden handelsübliche DVD-Videos von vorwiegend „High-Budget“ Filmen auf deren Dialoglevel und Dialnorm Metadaten-Wert hin untersucht. Während die nach  $L_{EQ(A)}$  gemessenen Dialoglevel stark variierten, war der Dialnorm-Wert kategorisch entweder beim Default-Wert von -27 oder beim Maximal-Wert von -31 gesetzt.<sup>29</sup> Daran wird unweigerlich erkennbar, dass sich das Verfahren zur Dialog-Normalisierung durch AC-3 nicht vernünftig durchsetzen konnte. Durch die Diskrepanz zwischen Dialnorm-Wert und tatsächlichem Dialoglevel funktioniert die abhörerseitige Dialog-Normalisierung nicht und zudem können störende Kompressionsartefakte auftreten, da die einzelnen Kennwerte der DRC relativ zum Dialog verschoben sind. Als problematisch muss erachtet werden, dass der Prozess des AC-3 Masterings bzw. DVD-Authorings meist durch andere Instanzen durchgeführt wird als die DVD-Mischung, sofern überhaupt eine Anpassung der ursprünglichen Kinomischung vorgenommen wird. Dadurch entsteht normalerweise kein Austausch zur sinnvollen Wahl des Dialnorm-Wertes; dieser bleibt üblicherweise unberührt auf dem Default-Wert -27 stehen oder wird zur Lautstärke-Maximierung auf -31 gesetzt. Eine optimale Umsetzung des sorgfältig durchdachten Standards scheitert hier abermals am menschlichen Fehler.

Bei der Produktion für TV und im ATSC Sendebetrieb wird von ähnlichen Problemen berichtet: Der Dialnorm-Wert sei meistens falsch gesetzt oder werde beim Default-Wert belassen, was auch auf die uneindeutige Bestimmung des Dialoglevels durch  $L_{EQ(A)}$  zurückgeführt wird. Als 2007 lediglich vereinzelte Sender digital nach ATSC sendeten, habe es sogar mehr Beschwerden über inkonsistente Lautstärken als bei NTSC Sendern gegeben.<sup>30</sup> Zusätzlich zur schlechten Umsetzung des AC-3 Standards wird unter Tonschaffenden zudem beklagt, dass bei einigen Sendern trotz Umstellung auf AC-3 und Dialog-Normalisierung die analoge Spitzenpegel-Limitierung nach älteren Aussteuerungs-Standards beibehalten wurde, was der Schaffung von mehr Dynamik widerspricht.

Ein besonderes Merkmal von AC-3, und möglicherweise auch Ursache für die auftretenden Missverständnisse, ist die doppelt freie Wahl des Lautstärke-Bezugs: Die Aussteuerung bei der Mischung und eben der Dialnorm-Wert beim Mastering können frei gewählt werden; beide sollen sich am Ende lediglich entsprechen. Feste

Vorgaben zur Aussteuerung und dementsprechend zur Dialnorm seitens US-amerikanischer Sender bekräftigen die Annahme, dass das Bedürfnis nach einer eindeutigen Aussteuerungs-Richtlinie für die Mischung besteht.

## **2.4. EBU R 128 und ATSC A/85 nach ITU-R BS.1770**

### **2.4.1. Differenzierung**

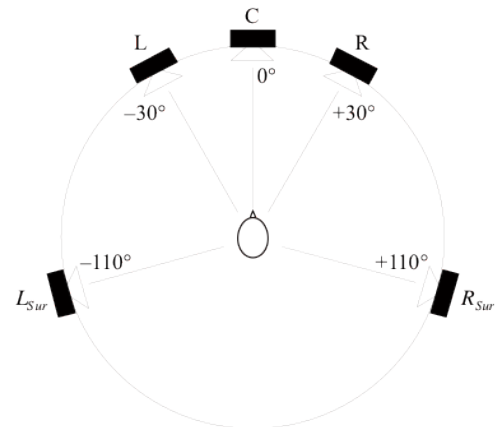
Die drei aufgeführten Dokumente werden gelegentlich verwechselt oder als austauschbar gehandhabt, weil sie in Teilen einander entsprechen. Sie haben jedoch jeweils eigenständige Funktionen und können wie folgt in Relation zu einander gesetzt werden: Das Dokument ITU-R BS.1770 (im Folgenden kurz BS.1770) und deren Erweiterungen BS.1770-1 bis BS.1770-3 definieren Algorithmen zur Messung von Programmlautheit und True Peak Level. Sie beschreiben also lediglich ein international einheitliches Aussteuerungs-Instrumentarium. EBU R 128 ist eine Empfehlung mit konkreten Richtlinien für die Aussteuerung im europäischen Rundfunk, die sich der in BS.1770 beschriebenen Algorithmen bedient, diese erweitert und rückwirkend beeinflusst hat. ATSC A/85 ist eine Empfehlung für den nordamerikanischen Raum sowie Mexiko, welche ebenfalls konkrete Richtlinien zur Aussteuerung definiert, die sich der BS.1770 Algorithmen bedienen, aber von der EBU R 128 abweichen.

### **2.4.2. Initiative**

Die Internationale Fernmeldeunion (ITU), Sonderorganisation der Vereinten Nationen, stellte einer Arbeitsgruppe im Jahr 2002 das Ziel, ein Lautstärke-Messgerät für TV und Hörfunk zu entwickeln, um Lautstärkesprünge zwischen Programmen und Sendern zu beseitigen.<sup>31</sup> Auf der Suche nach einem geeigneten Algorithmus, der bei typischen Rundfunkinhalten überzeugende Messungen liefert, wurden eingehende Versuche durchgeführt. Bei einer ersten Testreihe wurden weltweit an fünf Orten subjektive Datenbanken durch Hörtests aufgebaut, bei denen Testpersonen typische monophone TV- und Hörfunk-Inhalte ihrer Lautstärke nach an ein Referenzsignal anpassen sollten. Anschließend wurden zehn objektive Lautstärke-Messmethoden anhand dieser Datenbanken ausgewertet (viele davon waren anonyme Einsendungen von Entwicklerfirmen), bei der die  $L_{EQ}$ -Messung mit einer RLB Gewichtung (Revised Low-frequency B-curve, s. Abbildung 3) am besten abschnitt. Die RLB-Kurve ist eine leicht veränderte B-Kurve, die einen Hochpass bei 100Hz mit 6dB/Okt Flankensteilheit darstellt, und war in Kombination mit der  $L_{EQ}$ -

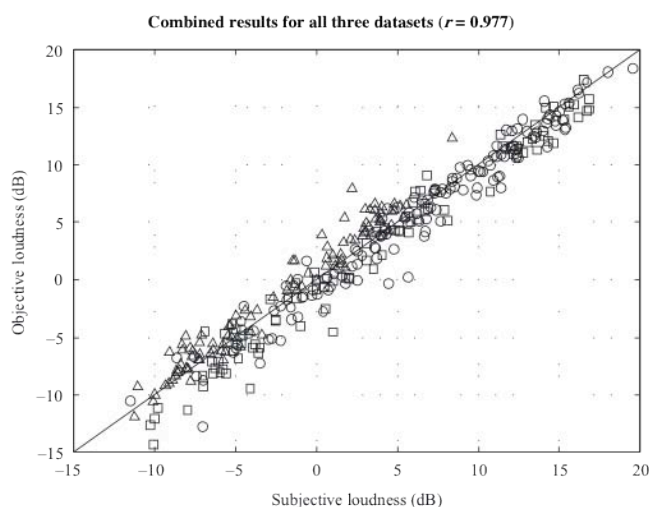


Messung vom Leiter der Studie Gilbert Soulodre selbst in die Testreihe eingebracht worden. Interessanterweise schnitt auch die ungewichtete  $L_{EQ}$ -Messung fast ebenso gut ab, woraus der Forscher schloss, dass die simplen energie-basierten Single-Band Messgeräte „ähnlich robust“ seien wie komplexere Varianten, die detaillierte Wahrnehmungs-Modelle enthielten. Die Ergebnisse stießen jedoch insgesamt auf Misstrauen<sup>32</sup> und so



**Abbildung 6:**  
Lautsprecheranordnung ITU-R BS.775

sollten zwei weitere Testreihen mit unterschiedlicheren Testsignalen und auch Stereo- sowie Mehrkanalinhalten durchgeführt werden. Bei diesen Untersuchungen kam die 5.1 Lautsprecheranordnung ITU-R BS.775 zur Anwendung (Abbildung 6). Mehrkanalige Klangbeispiele sollten vor allem im Hinblick auf die Nutzung der Surrounds verschiedene Mischungsstile enthalten und monophones Testmaterial sollte teilweise über den Center-Lautsprecher sowie teilweise als „dual mono“ über Rechts und Links wiedergegeben werden. Nach Abschluss und Auswertung dieser Testreihen veröffentlichte die ITU 2006 dann erstmals eine Empfehlung BS.1770 für Algorithmen zur Messung von Programmlautheit und True-Peak Level im Rundfunk. Der entwickelte Lautheits-Algorithmus war mit einer Korrelation von  $r = 0,977$  (Abbildung 7) durch die subjektiven Daten der Testpersonen validiert worden und wird seitens der ITU ausdrücklich auch aufgrund seines simplen Aufbaus und minimaler Rechenanforderung als besonders geeignet für die Rundfunk-Praxis bezeichnet.<sup>33</sup>



**Abbildung 7:**  
Korrelation subj. Wahrnehmung zu ITU-R BS.1770-1

### 2.4.3. Verfahren

Bei dem ITU Lautheits-Algorithmus durchlaufen alle Kanäle eines Mehrkanal-Audiosignals einen identischen Prozess. Lediglich der LFE-Kanal bleibt derzeit unberücksichtigt; laut EBU finden aber bereits Untersuchungen zum Einfluss des LFE auf die Lautheit eines Programms statt. Die

ursprüngliche RLB-Gewichtung wird jeweils durch einen identischen Pre-Filter erweitert, einer Highshelf-Anhebung von 4dB ab ca. 1kHz, die dem akustischen Einfluss des Kopfes bei der Schallwahrnehmung Rechnung tragen soll. Die daraus resultierende Gewichtung wurde zunächst R2LB, also zweifach überarbeitete tief-frequente B-Gewichtung, und inzwischen K-Gewichtung genannt (Abbildung 8). Sie steht in keinerlei Verbindung zum

K-System von Bob Katz, das sich seinerseits verschiedener Lautheits-Algorithmen, unter anderem dem aus BS.1770 bedient. Anschließend wird jeweils vom gefilterten Signal  $y$  des jeweiligen Kanals  $i$  ( $i \in I$  wobei  $I = \{L,C,R,Ls,Rs\}$ ) eine RMS-Messung über ein Zeitintervall  $T$  sowie eine Kanal-spezifische Gewichtung  $G_i$  durchgeführt, weil davon

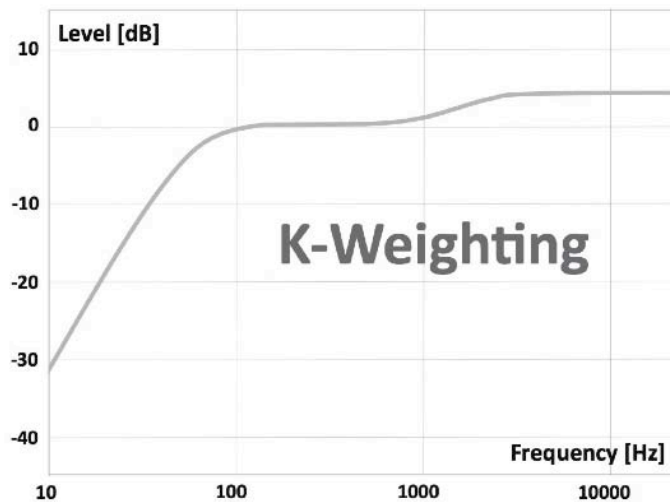


Abbildung 8: K-Gewichtung

ausgegangen wird, dass Signale hinter dem Hörer lauter wahrgenommen werden als vor ihm. Die Einzelwerte werden anschließend zu einer Gesamtlautheit  $L_K$  summiert, die logarithmisch dargestellt und anschließend so normalisiert wird, dass ein Sinuston von  $0 \text{ dB}_{FS}$  auf einem der Frontkanäle einer Gesamtlautheit von  $-3,01 \text{ LKFS}$  entspricht:

$$L_K = -0.691 + 10 \log_{10} \sum_i G_i \cdot z_i \quad \text{wobei} \quad z_i = \frac{1}{T} \int_0^T y_i^2 dt \quad (\text{RMS über Zeit } T)^{34}$$

Die Lautheit wird angegeben in LKFS („Loudness, K-Weighting, referenced to digital Full Scale“). Die EBU bezeichnet die gleiche Einheit als LUFS („Loudness Units, referenced to digital Full Scale“) um eine Verwechslung mit dem genannten K-System von Bob Katz zu vermeiden. Diese zweite Bezeichnung entspricht dem internationalen ISO Namen-Standard 80000-8; eine zukünftige Einigung auf LUFS scheint also naheliegend. In dieser Arbeit wird für jegliche Messungen nach BS.1770 und A/85 die Einheit LKFS sowie für Messungen nach R 128 die Einheit LUFS verwendet, wengleich diese den Algorithmen aus BS.1770-2 entsprechen. LU („Loudness Units“) ist die Einheit der relativen Lautheit und ist äquivalent zu dB.<sup>35</sup>

- ***Programme Loudness***

Besetzt man das bisher unbekanntes Zeitintervall  $T$  durch die gesamte Dauer eines Programms, dann erhält man die Programmlautheit. Alternativ dazu kann natürlich über benutzerdefinierte Zeitabschnitte eines Programms gemessen werden („Integrated Loudness“).

Die Projektgruppe P/LOUD der European Broadcast Union (EBU), bestehend aus Rundfunk Vertretern sowie Soft- und Hardware Entwicklern, beschäftigte sich seit 2008 ebenfalls eingehend mit der Thematik und untersuchte Möglichkeiten, leise oder stille Anteile einer Mischung aus der Messung auszuschließen, um unnötig laute Übertragungen von solchen Programmen zu verhindern. Der Vorschlag eines absoluten Gates, welches alle Anteile unterhalb von  $-70$  LUFS aus der Messung ausschließt in Kombination mit einem relativen Gate, das  $8$  LU unter der absolut gegateten Programmlautheit liegt, wurde von der ITU aufgegriffen und ab BS.1770-2 mit einer leichten Absenkung auf  $10$  LU übernommen.<sup>36</sup>

- ***Momentary und Short-Term Loudness***

Die EBU definierte noch zwei weitere feste Zeitintervalle: Eine Momentary Loudness soll demnach durch Integration mit einer Fensterung von  $400$  ms, eine Short-Term Loudness durch Integration mit einer Fensterung von  $3$  s berechnet werden. Beiden Größen liegt eine ungegatete Lautheit nach BS.1770 zugrunde. Sie ermöglichen ein Aussteuern nach Lautheit in Echtzeit, anders als bei einer rückwirkenden Anpassung der Programmlautheit nach abgeschlossener Mischung.<sup>37</sup>

- ***Loudness Range***

Ebenfalls von der EBU wurde ein Algorithmus der Firma TC Electronic zur Bestimmung der Lautstärke-Dynamik eines Programms konstituiert. Die Loudness Range wird berechnet, indem zunächst die Lautheit nach BS.1770 über ein gleitendes Analyse-Fenster von  $3$  s integriert wird. Aufeinander folgende Fenster sollen sich dabei um mindestens  $66\%$ , also  $2$  s überlappen. Anschließend greift erneut ein absolutes Gate bei  $-70$  LUFS sowie ein relatives Gate bei diesmal  $20$  LU unterhalb der absolut gegateten Lautheit um Passagen zu ignorieren, welche z.B. ausschließlich Hintergrund Geräusche wie Atmosphären enthalten. Anschließend wird die Quantität der einzelnen Lautheitswerte ausgewertet und lediglich der Lautheitsunterschied zwischen dem  $10.$  Und  $95.$  Perzentil als

Loudness Range in LU definiert. Dieser letzte Schritt soll verhindern, dass sehr kurze laute Passagen oder das Ausblenden von Musik in die Berechnung mit einfließen.<sup>38</sup>

- **True Peak Level**

Beim True Peak Level nach BS.1770 handelt es sich nicht um eine Lautheitsgröße, dennoch spielt sie in allen drei erwähnten Empfehlungen eine zentrale Rolle, da sie die technische Intersample-genaue Aussteuerung des Audio-Signals ermöglicht und damit Vorgänger wie VU und QPPM ablöst. Abbildung 9 zeigt den Signalweg durch eine 12,04 dB Absenkung (Attenuation) zu Headroom Zwecken, eines vierfachen Oversamplings zur Bestimmung von Intersample-Peaks, einer optionalen Prä-Emphasis ab 14,1 kHz zur Berücksichtigung möglicher Verfälschungen an Aliasing Filtern und einer optionalen DC Eliminierung. Sowohl eine absolute Ausgabe des True Peak Levels, als auch der für Sender-Konformität wichtige Maximalwert in dBTP sind in Messgeräten vorgesehen.<sup>39</sup>

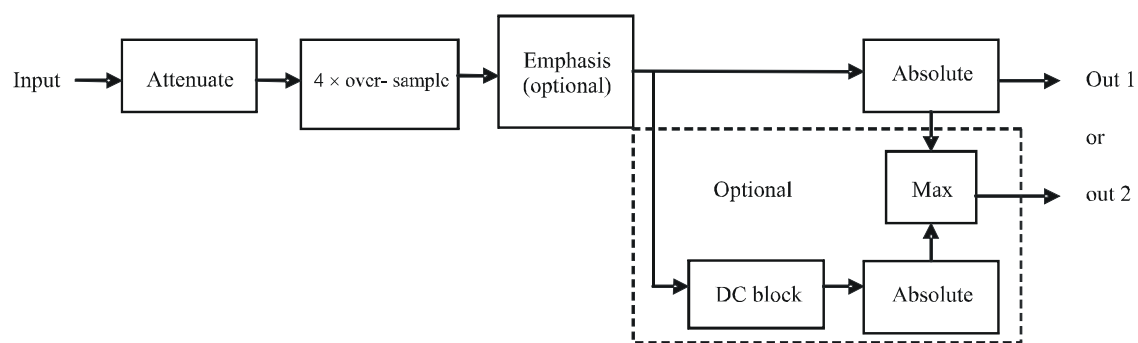


Abbildung 9: Ermittlung des True Peak Level

#### 2.4.4. Umsetzung

- **ATSC A/85**

2009 reagierte die ATSC auf die Untersuchungen vom ITU-R mit der Empfehlung A/85, welche Lautheits-Messungen von Dialogpassagen (nicht des ganzen Programms) nach ITU-R BS.1770-1 vorsieht und somit  $L_{EQ(A)}$  ablöst (Kapitel 2.3.3.). Die Gating-Funktion von BS.1770-2 ist in dem ATSC Dokument also noch nicht berücksichtigt. Sollte das Programm keinen Dialog beinhalten, dürfen Passagen eines alternativen Anker-Elements (Ein Element, auf das sich ein durchschnittlicher Zuschauer konzentrieren würde) gemessen werden. Erst wenn es nicht möglich oder „nicht praktisch“ ist, die Messung auf ein Anker-Element wie Dialog zu fokussieren, soll eine Programmlautheits-Messung durchgeführt

werden, wobei aufgrund des Fehlens eines Gates die Messung bei besonders leisen Passagen oder Stille unterbrochen werden soll. Dabei ist es prinzipiell gleich, wo dieser gemessene Wert liegt, solange der Dialnorm Metadaten-Wert von AC3 diesem entspricht. Besonders im Hinblick auf den Datentransfer ohne Metadaten oder wenn seitens der beteiligten Parteien (Sender) keine anderen Vorgaben gemacht werden (s. Vorgaben PBS, NBC, Discovery, Kapitel 2.3.3.), soll ein Wert von -24 LKFS angestrebt werden. Soft- bzw. Hardware-bedingte Schwankungen von  $\pm 2$  LU sind akzeptabel, True Peak Level sollen unter -2 dBTP gehalten werden um problemlose Downmixe zu ermöglichen.<sup>40</sup> Die Empfehlung stellt daher gewissermaßen einen Kompromiss zwischen den zwei Standards AC-3 und BS.1770 zur Verbesserung der Lautstärke-Kontinuität im Rundfunk dar. Am 13. Dezember 2012 trat außerdem ein weiteres Gesetz des FCC (Federal Communications Commission) in Kraft, nachdem im US-Kongress der „CALM Act“ (Commercial Advertisement Loudness Mitigation Act) verabschiedet wurde, der gesetzlich vorschreibt, dass TV-Werbungen die Vorgaben der A/85 befolgen müssen.<sup>41</sup>

- ***EBU R 128***

Im August 2010 veröffentlichte Die European Broadcast Union die von der P/LOUD Arbeitsgruppe entworfene Empfehlung EBU R 128 für den europäischen Rundfunk. Das kurz gefasste Haupt-Dokument empfiehlt, dass Programme in ihrer Gesamtlänge durch ein Lautheits-Messgerät mit Gatefunktion (EBU Mode) nach BS.1770-2 gemessen und auf eine Programmlautheit von -23LUFS ( $\pm 1$  LU Toleranz) normalisiert werden. Ausdrücklich soll sich die Messung nicht auf bestimmte Elemente wie Dialog, Musik oder Soundeffekte beschränken, sondern sich über das gesamte Programm erstrecken. Des Weiteren sollen Programme ein maximales True Peak Level von -1 dBTP nicht überschreiten und sollten zusätzlich zur Programmlautheit auch durch ihre Loudness Range beschrieben werden. In sog. „Tech Docs“ werden zudem die Verfahren zur Generierung der einzelnen Werte erläutert. Die EBU stellt der nordamerikanischen Dialog-Normalisierung eine Lautheits-Normalisierung gegenüber.

Zwei Jahre später wird die Empfehlung bereits von einigen europäischen, darunter vielen öffentlich-rechtlichen Sendern umgesetzt. Einzelne Sender wie ARTE haben die Empfehlung zudem durch Vorgaben zur maximalen Loudness Range (20 LU), zur maximalen Short-Term Loudness bei kurzen Programmen

unter zwei Minuten (-20 LUFS) sowie zur Loudness Range von Dialogpassagen erweitert (-23 LUFS  $\pm$ 7 LU).<sup>42</sup>

Die P/LOUD Gruppe begründet die Wahl von -23 LUFS mit Untersuchungen zwischen ungegateden Lautheits-Messungen nach A/85 (Zielwert bei -24 LKFS) und gateden Messungen nach ITU-R BS.1770-2, bei denen sie zu dem Schluss kamen, dass sich beide am ehesten entsprechen, wenn letztere bei -23 LUFS liegen.<sup>43</sup>

Eine besondere Betrachtung verdient bei der Lautheits-Normalisierung der Dialog, da man annehmen könnte, dass dieser etwa in Programmen mit lauten Musikpassagen und Action-Sequenzen leiser sowie in Programmen mit ausschließlich Dialog und leisen Hintergrundgeräuschen lauter ausfallen könnte. Prinzipiell lässt sich aber festhalten, dass bereits bei der Mischung der Dialog durchschnittlich bei -23 LUFS gehalten werden soll. Die Mischung ist also demnach nicht einfach frei und wird nach Abschluss normalisiert. Für Programme mit vielen lauten Passagen bedeutet das natürlich, dass hier die Dynamik eingeeignet werden muss bzw. diese Elemente nicht unverhältnismäßig viel lauter als der Dialog sein dürfen, um sowohl eine Programmlautheit als auch einen Dialog im Bereich von -23 LUFS zu erreichen.

### 2.4.5. Erfolg

Betrachtet man die Auswirkungen der Aussteuerung nach BS.1770 im Rundfunk, muss unweigerlich der Blick auf die Dynamik gerichtet werden. Bei der bisherigen Verfahrensweise einer Pegelung nach technischer Vollaussteuerung, resultierten bislang deutlich geringere Lautstärken für Programme mit großer Dynamik (z.B. Klassikkonzerte, Filme) als für Programme mit besonders kleiner Dynamik (z.B. Werbung). Eine solche Peak-Normalisierung, also die Gleichschaltung von Sendematerial nach einem technischen Pegel stört also nicht nur Zuschauer wegen enormer Lautstärke-Schwankungen, sie provoziert auch eine unnatürlich starke Dynamikeinengung

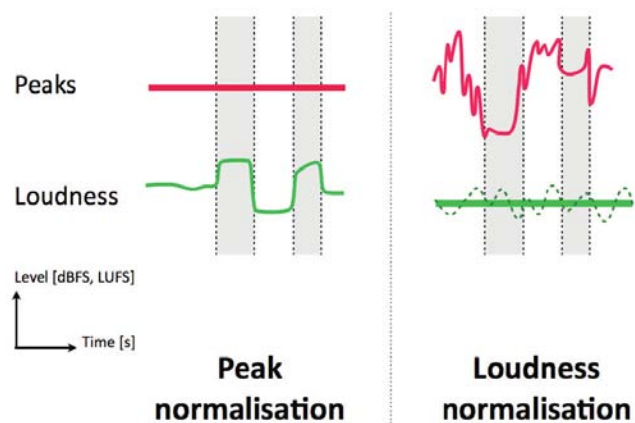


Abbildung 10: Schematische Darstellung von Peak- und Lautheits-Normalisation

durch Kompressoren und Limiter bei der Produktion und somit einen zusätzlichen Verlust an Klangqualität. Der sog. „Loudness War“ wird durch eine Lautheits-Normalisierung entschärft, da eine künstliche Steigerung der Lautstärke durch Überkompression und somit ein Konkurrenz-Vorteil bei der Sendung nicht mehr erwartet werden kann (Ausnahme s. Kapitel 2.2.4.).

Für Erhebungen zur Zufriedenheit der Zuschauer ist es derzeit noch sehr früh. Von den Tonschaffenden wurde die Umstellung auf Lautheits-Normalisierung aber sehr unterstützt und begrüßt. Anders als in der Vergangenheit müssen sich Mischtonmeister nicht durch ihre Fähigkeit profilieren in einem technisch vorgegebenen Rahmen laute Mischungen zu erzeugen, sondern können sich darauf konzentrieren, einen dramaturgisch sinnvollen Lautstärke-Verlauf zu erarbeiten, der sowohl zum Programm als auch zu den Abhörbedingungen der Konsumenten passt und sie dürfen dabei insgesamt mehr Dynamik zulassen. Die Veränderungen bringen also auch einen politischen Aspekt mit sich: Sie verändern die Erwartungen von Produzenten an Mischtonmeister im Hinblick auf die Qualität ihrer Arbeit und befreien diese gewissermaßen von der Karte des „Schwarzen Peter“.

Das lässt sich ebenso auf Perspektiven für die Kinomischung übertragen, denn bei dieser bewältigen Mischtonmeister derzeit regelmäßig eine Gratwanderung zwischen der Einhaltung eines Standards und Produzenten-Interessen an einer lauten Mischung, wobei eine Missachtung der letzteren ihnen die Folgeaufträge kosten kann. Nähme man den Aspekt der Lautstärke-Maximierung aus der Gleichung, würden also ähnlich wie jetzt im Rundfunkbereich Konflikte im eigentlich kreativen Prozess der Mischung vermieden werden.

---

<sup>1</sup> vgl. [1] Skovenborg & Nielsen: Evaluation of Different Loudness Models with Music and Speech Material, S.2

<sup>2</sup> vgl. ebenda, S. 5f

































<sup>3</sup> siehe auch [2] Zwicker & Fastl: Psychoacoustics - Facts and Models, S. 64

<sup>4</sup> vgl. [1] Skovenborg & Nielsen: Evaluation of Different Loudness Models with Music and Speech Material, S.6

<sup>5</sup> für eingehende Untersuchungen zu Lautheits-Messgeräten sei empfohlen [3] Naumann, Silvio: Entwicklung eines Verfahrens zur Generierung, Übertragung und Auswertung lautstärkerelevanter Kennwerte

<sup>6</sup> im Folgenden vgl. [4] Allen, Ioan (Dolby Laboratories): Are Movies too loud?

<sup>7</sup> „Der ‚Ride‘-Film ist inzwischen Normalität in Hollywood und wenngleich der Film spannend sein kann, hat die Bezeichnung ‚Schnitt-zur-Verfolgungsjagd‘ keine Bedeutung mehr - Die Verfolgungsjagd beginnt regelmäßig am Anfang von Rolle 1!“

- 
- <sup>8</sup> vgl.  TASA: The TASA Standard, S. 7
- <sup>9</sup> vgl.  Allen, Ioan (Dolby Laboratories): Are Movies too loud?
- <sup>10</sup> vgl.  Skovenborg & Nielsen: Evaluation of Different Loudness Models with Music and Speech Material, S.25
- <sup>11</sup> vgl.  TASA History,  Who are we?
- <sup>12</sup> vgl.  FSK: Aufgaben und Strukturen
- <sup>13</sup>  ISO 21727:2004, Cinematography - Method of measurement of perceived loudness of motion-picture audio material
- <sup>14</sup> „(...) Trailer zu mischen ist nicht mischen sondern verzerren...“
- <sup>15</sup> vgl.  DUC: Dolby 7. Is everybody mixing at this level? (matt-o-, 14.05.11, 06:05)
- <sup>16</sup> vgl.  DUC: DCP headroom?/levels? (matt-o-, 09.06.12, 02:56)
- <sup>17</sup>  Moore, Glasberg, Stone: Why Are Commercials so Loud? - Perception and Modeling of the Loudness of Amplitude-Compressed Speech
- <sup>18</sup> zitiert in  Skovenborg & Nielsen: Evaluation of Different Loudness Models with Music and Speech Material, S.6
- <sup>19</sup> siehe  Riedmiller, Lyman, Robinson (Dolby Laboratories): Intelligent Program Loudness Measurement and Control - What Satisfies Listeners? S. 5f
- <sup>20</sup>  Henle, Hubert: Das Tonstudio Handbuch, S. 101-105
- <sup>21</sup> vgl.  Benjamin, Eric: Preferred Listening Levels and Acceptance Windows for Dialog Reproduction in the Domestic Environment
- <sup>22</sup> für eingehende Untersuchungen zum Thema AC-3 sei empfohlen  Ziegler, David: Der AC-3 Standard in Theorie und Praxis
- <sup>23</sup> vgl.  Vinton & Robinson (Dolby Laboratories): Automated Speech/Other Discrimination for Loudness Monitoring
- <sup>24</sup> Download unter <http://www.dolby.com/us/en/professional/technology/broadcast/dialogue-intelligence.aspx>
- <sup>25</sup>  Henle, Hubert: Das Tonstudio Handbuch, S. 102
- <sup>26</sup> vgl.  ATSC: A/53 Part 5 - AC3 Audio System Characteristics, S. 8f
- <sup>27</sup> vgl.  Dolby: Discovery's Deadliest Catch Customer Story
- <sup>28</sup> vgl.  Media Channel: Sendungen in 5.1 Dolby Digital Surround Sound - TV und Hörfunk
- <sup>29</sup> vgl.  Ziegler, David: Der AC-3 Standard in Theorie und Praxis, S. 47-53
- <sup>30</sup> vgl.  Jacobs, Bruce (Broadcast Engineering): Dialnorm - A good idea gone bad?
- <sup>31</sup> vgl.  Soulodre, G.A.: Evaluation of Objective Loudness Meters
- <sup>32</sup> vgl.  Spikofski (IRT): Lautstärkemessung im Rundfunk - Stand der internationalen Standardisierung, S. 17
- <sup>33</sup> vgl.  ITU-R: Recommendation ITU-R BS.1770-3, S.7-16
- <sup>34</sup> vgl. ebenda, S.5
- <sup>35</sup> vgl.  EBU: TECH 3343 - Practical guidelines EBU R 128, 9f
- <sup>36</sup> vgl. ebenda, S. 11
- <sup>37</sup> vgl. ebenda, S. 6
- <sup>38</sup> vgl.  EBU: TECH 3342 - Loudness Range EBU R 128, S. 6
- <sup>39</sup> vgl.  ITU-R: Recommendation ITU-R BS.1770-3, S.17f
- <sup>40</sup> vgl.  ATSC: Recommended Practice A/85, S. 17ff
- <sup>41</sup> vgl.  Moulton, Dave (TVTechnology): LKFS & The CALM Act
- <sup>42</sup> siehe  Burg, Lothaire (ARTE): Memo für Produzenten & Dienstleister - Lautheit von Fernsehprogrammen
- <sup>43</sup> vgl.  EBU: TECH 3343 - Practical guidelines EBU R 128, 11f



## 3. Praktische Lautheitsmessungen bei Kinofilmen

### 3.1. Auswahl der Algorithmen und zu erfassenden Kennwerte

Die im 2. Kapitel untersuchten Methoden zur Lautheits- bzw. Dialog-Normalisierung, die bei Kinowerbungen und -trailern, Heimkino-Formaten und beim Rundfunk Anwendung finden, werden im Folgenden auf ihre Aussagekraft im Hinblick auf die Lautstärke von Filmen und eine Modellnorm im Kino geprüft:

1. Programmlautheit nach  $L_{EQ(M)}$
2. Programmlautheit nach  $L_{EQ(A)}$
3. Programmlautheit nach ITU-R BS.1770-1 (ohne Gate)
4. Programmlautheit nach ITU-R BS.1770-2 (mit Gate)
5. Dialoglevel nach  $L_{EQ(A)}$  mit Dialogue Intelligence
6. Dialoglevel nach ITU-R BS.1770-1 (ohne Gate) mit Dialogue Intelligence

(Eine Dialoglevel-Messung nach ITU-R BS.1770-2 (mit Gate) schließt sich aus, da Dialogue Intelligence selbst eine Gate-Funktion darstellt)

Darüber hinaus sollen Werte der entwickelten Algorithmen in Betracht gezogen werden, die zwar nicht unmittelbar Grundlage zur Normalisierung sind, jedoch die Audio-Signale weiter beschreiben:

7. Short-Term Loudness nach ITU-R BS.1770-2 (maximal und zeitabhängig)
8. Momentary Loudness nach ITU-R BS.1770-2 (maximal und zeitabhängig)
9. True Peak Level nach ITU-R BS.1770-2 (maximal und zeitabhängig)
10. Loudness Range nach ITU-R BS.1770-2
11. Sprachanteil ermittelt durch Dialogue Intelligence

### 3.2. Erstellung der Film-Datenbank

Die Messungen sollten möglichst an Filmen durchgeführt werden, die ein Kinopublikum erreicht haben und auf 35mm mit Dolby SRD erschienen sind, um sicherzustellen, dass sie ein Printmastering, also die Prüfung durch einen Dolby Consultant durchlaufen haben. Sie sollten sich zudem aus verschiedenen Genres und möglichst aus Produktionen verschiedener Länder zusammensetzen, sowie in den letzten zehn Jahren veröffentlicht worden sein. Dadurch sollten die Messungen eine Vielzahl von Funktionen erfüllen:

- Eine Übersicht zur aktuellen Lautstärke-Situation im vom Dolby Standard geprägten Kinosystem verschaffen,
- nationale Aussteuerungs-Tendenzen aufzeigen,
- Einflüsse von Genres untersuchen,
- Film-übergreifende Kontinuität des Dialoglevels darstellen,
- eine Film-Datenbank mit zugehörigen Lautheits-Kennwerten bilden, die in Verbindung mit einer subjektiven Bewertung der Filme Aufschluss über die Tauglichkeit der einzelnen Algorithmen und Kennwerte für eine Lautheits-Normalisierung im Kino gibt.

Zur Beschaffung der Filme zu Messungs-Zwecken wurden 70 Verleiher kontaktiert, von denen sieben Interesse an dem Projekt zeigten und mit zweien eine Zusammenarbeit zustande kam. Constantin Film Verleih GmbH stellte der Studie in Kooperation mit ARRI München 27 Filme, Prokino Filmverleih GmbH 13 Filme zur Verfügung, von denen jeweils die 5.1 Printmaster Kinomischungen ohne Bild aus DA88, DVD-R und Festplatten aus dem Backup restoret wurden. Die Messungen fanden im Oktober 2012 in München statt.

In einer zweiten Phase stellten das Tonstudio Hanse Warns sowie der französische Mischtonmeister Bruno Tarrière weitere Printmaster zur Verfügung. Die Messungen fanden im November und Dezember 2012 in Berlin statt.

Insgesamt wurden 50 Filme gemessen, davon 43 Originalversionen und sieben Synchronfassungen. 22 Filme wurden in Deutschland gemischt, zehn in den USA oder von US-amerikanischen Mischtonmeistern, sieben in Frankreich und sechs in England. Etwa 19 Filme können am besten als Drama bezeichnet werden, zwölf als Komödie, elf als Thriller bzw. Action oder Horror, drei als Musikfilm. Einige Filme wurden von gleichen Mischtonmeister gemischt, darunter fünf von Matthias Lempert, vier von Michael Kranz mit Benjamin A. Rosenkind, drei von Martin Steyer, drei von Bruno Tarrière, drei von Changis Chahrokh, drei von Leslie Shatz, sowie jeweils zwei von Vincent Arnardi und Bob Beemer.

Allerdings ist eine eindeutige Tendenz zu internationalen Koproduktionen festzustellen; bis zu sechs beteiligte Nationen sind dabei keine Seltenheit. Noch dazu fand die Postproduktion inklusive Mischung nicht selten in anderen Ländern als dem des Regisseurs bzw. der Hauptdrehorte statt. Es gab auch Konstellationen, bei denen Mischtonmeister Filme im Ausland mischten. Rückschlüsse auf nationale Tendenzen sollten also, auch aufgrund der ohnehin eingeschränkten Datenlage, nur unter Vorbehalt gemacht werden.

Die Filme wurden von verschiedenen Backup Datenträgern umgespielt, welche die Audio-Master zwar allesamt digital speichern, jedoch durch Umspielprozesse (z.B. durch analoge Zwischenstufen) nicht zwangsläufig eine exakte Rekonstruktion des Originals garantieren können. Wenngleich dieser Faktor zunächst unwichtig erscheint, zumal ein Teil der Filme ohnehin analog gemischt wurde, sind jedoch bei einigen Mastern keine Referenzpegel und/oder keine eindeutigen Beschriftungen vorhanden, wodurch eine Originaltreue zwar angenommen, jedoch nicht bewiesen werden kann. Zwei Fehler können deshalb in einzelnen Fällen nicht ausgeschlossen werden:

1. Es sind keine Referenzpegel vorhanden, wodurch der Messung falsche absolute Pegel zugrunde liegen könnten. Möglich ist dieser Fehler bei Umspielung von DA88, wenngleich vorhandene Referenzpegel gewöhnlich stimmten oder durch den Autor korrigiert wurden (in drei Fällen bei Abweichung von  $\pm 2$  dB). Die zwei Filme mit nicht vorhandenen Pegeln sind in der Datenbank entsprechend markiert. Zusätzlich wird aufgeführt, von welcher Art Datenträger die Audiomaster stammen.
2. Es liegt keine eindeutige Beschriftung zur Art der Mischung vor. Generell weisen Beschriftungen wie PM für Print Master, SRD für Dolby SR mit Digital, DMU für die Dolby Mastering Unit, DCP für Digital Cinema Package oder z.B. das Vorliegen eines 5.1 Tracks in Verbindung mit einem SVA Track darauf hin, dass es sich eindeutig um eine Kinomischung handelt und nicht etwa um eine DVD-Mischung. 40 von 49 Filmen hatten eine eindeutige Kennzeichnung und tragen in der Datenbank eine zusätzliche garantierende Kennzeichnung „PM“. Bei den übrigen 9 Filmen kann das Vorliegen einer Kinomischung zwar angenommen, jedoch nicht garantiert werden.

Zwar wäre es problematisch, bei solchen Filmen, die evtl. falsch umgespielt wurden oder als DVD-Fassung vorliegen, Rückschlüsse auf Mischungsstile oder die aktuelle Lautstärke-Situation im Kino zu ziehen, zum Evaluieren einer Mischungsnorm sind sie jedoch gleichermaßen geeignet. Eine DVD-Fassung z.B., die durchschnittlich womöglich lauter aber weniger dynamisch gemischt wurde, muss auch technisch als lauter erkannt werden, wenngleich sie nach entsprechender Absenkung weiterhin eine kleinere Dynamik als eine Kinomischung haben wird.

### 3.3. Subjektive Einschätzungen

Um die Kennwerte der Datenbank in Relation zur subjektiven Wahrnehmung zu setzen und auf ihre Aussagekraft hin zu überprüfen, wurden vom Autor auch Einschätzungen zur optimalen Wiedergabelautstärke der einzelnen Filme getroffen. Diese „subjektive Pegelung“, wie sie z.B. auch gewöhnlich vom Filmvorführer zu Beginn eines Screenings vorgenommen wird, weist jedem der gemessenen Filme einen Wert in dB zu, der die Abweichung der angestrebten Wiedergabelautstärke von Dolby Level 7 beschreibt. Sie kann alternativ also auch als gewünschtes Wiedergabelevel auf der Dolby Skala zwischen 1 und 10 dargestellt werden (z.B. subjektive Pegelung von -5 dB: der Film sollte auf Dolby Level 5,5 wiedergegeben werden<sup>1</sup>). Bei der Erfassung dieser Daten ging es weniger darum, absolut gültige Werte für die Wiedergabe dieser Filme vor einem Kinopublikum zu finden, sondern vielmehr um die Frage, wie eine individuelle Person die Filme relativ zueinander wiedergeben würde, natürlich in Bezug zur persönlichen Ideallautstärke. Wenngleich sich die Ideallautstärke eines Einzelnen nicht auf andere Zuschauer übertragen lässt, ist zumindest bei den relativen Verhältnissen von Film zu Film von einer gewissen Allgemeingültigkeit auszugehen.<sup>2</sup> Ideal wäre eine solche Studie mit zahlreichen unabhängig voneinander getesteten Personen gewesen um willkürliche Abweichungen einzudämmen und nach einer Normierung der jeweiligen Ideallautstärke gefestigtere Ergebnisse zu erzielen. Die Audiomaster durften jedoch die Postproduktionshäuser nicht verlassen, was ein solches Vorhaben ausschloss. Für diese Arbeit und die bloße Demonstration eines Modells zur Lautstärke-Normalisierung im Kino wird die beschriebene Annäherung als ausreichend erachtet.

Zusätzlich zur subjektiven Pegelung wurden anschließend für einige Filme Anmerkungen bei besonders lauten Stellen gemacht und wie laut diese Stellen - ebenfalls im Verhältnis zu Dolby Level 7 - maximal wiedergegeben werden sollten. Durch diese Vorgehensweise sollte verhindert werden, dass Filme lediglich anhand ihrer lautesten Passagen eingepegelt würden.

### **3.4. Durchführung der Messungen**

Um möglichst repräsentative Ergebnisse für die Datenbank zu generieren, sollten unnötige Wandlungen vermieden werden, weshalb rein File-basierte Messungen mit Software-Lösungen sinnvoll erschienen, die sich in ihrem Segment am Markt bereits bewährt und etabliert hatten. Die einzelnen Rollen wurden aneinandergesetzt, auf Bildstart bei TC 10:00:00:00 getrimmt um Referenzpegel sowie Startpieps auszuschließen, als 5.1-Monofiles konsolidiert und anschließend File-basiert von zwei Programmen gemessen:

#### **3.4.1. NUGEN LMB**

Der NUGEN Loudness Management Batch Processor stellt die Standalone Stapelbearbeitungs-Version von den weit verbreiteten NUGEN VisLM-C und VisLM-H dar. Durch seine Struktur ist er auf Finalisierungen bzw. Massen-Anpassungen an Lautheits-Vorgaben wie EBU R 128 oder ATSC A/85 spezialisiert, wie sie in größeren Postproduktions-Häusern oder bei Sendern auftreten. Dateien werden aus einem „Watched Folder“ geladen und nach einem definierten Algorithmus (z.B. ITU-R BS.1770-2) gemessen. Anschließend werden entweder lediglich die Werte für Integrated Loudness, Max. Short-Term Loudness, Max. Momentary Loudness, Max. True Peak und Loudness Range ausgegeben oder eine Korrektur nach Vorgaben derselben vorgenommen.

Die Verläufe der zeitlichen Kennwerte können dabei in einem Graph nachempfunden werden und gehen zusätzlich aus einer vom LMB erzeugten log-Datei hervor. Gemessen werden kann aus einer Vielzahl von Datei-Formaten, unter anderem .MXF, was als Audio-Hüllformat für DCPs dient. Das macht es theoretisch möglich, Kinomischungen unmittelbar anhand der DCPs zu messen, allerdings verhindert die KDM (Key Delivery Message) bei encrypteden kommerziellen Kopien einen direkten Zugriff auf die Dateien. NUGEN LMB wurde zur Bestimmung von Programmlautheit, Short-Term Loudness, Momentary Loudness, True Peak Level und Loudness Range nach ITU-R BS.1770-2 verwendet.

Am 23. Januar 2013, also nach Abschluss der Messungen, brachte NUGEN ein Update für VisLM mit integriertem Dolby Dialogue Intelligence Gate heraus. Ein vergleichbares Update für LMB war zu diesem Zeitpunkt nicht in Planung.

### 3.4.2. *Dolby Media Meter 2*

Das DMM2 stellt die Software-Variante zu Dolbys Program Optimizer DP600 dar. Wenngleich es sich für Lautheits-Messungen nach R 128 weniger durchsetzen konnte, führt es File-basierte Messungen bis zu 7.1 nach  $L_{EQ(A)}$  sowie ITU-R BS.1770-1 und -2 durch und zwar jeweils mit oder ohne Dialogue Intelligence (1770-2 nur ohne DI). Auch Messungen von bereits enkodiertem Material wie Dolby Digital sind möglich, womit das Programm besonders auf den hauseigenen AC3-Standard und vor allem natürlich auf die Ermittlung des wichtigen Dialnorm-Parameters ausgerichtet ist. Zusätzlich wird unabhängig vom benutzten Algorithmus auch der ebenfalls von Dolby entwickelte  $L_{EQ(M)}$ -Wert des Audio-Materials sowie Sample Peak oder True Peak ausgegeben. Beim DMM2 ist nach abgeschlossener Messung keine Korrektur vorgesehen. Es spielt jedoch ebenfalls eine log-Datei und einen Graphen aus, der auch Dialogstellen im Film kennzeichnet und anders als der LMB die zeitliche Entwicklung von Integrierter Lautheit und Integriertem Dialoglevel visualisiert. Das DMM2 wurde zur Bestimmung von Programmlautheit sowie Dialoglevel nach  $L_{EQ(A)}$  und ITU-R BS.1770-1 sowie Programmlautheit nach  $L_{EQ(M)}$  und Sprachanteil der Filme verwendet. Bei  $L_{EQ(M)}$ -Werten unterhalb von 72 wird vom DMM2 kein Wert, sondern nur die Bemerkung „Low“ ausgegeben. Zur Berechnung dieser Werte wurden deshalb zusätzliche Messungen mit zuvor verstärktem Eingangssignal durchgeführt.

Originaltitel	Deutscher Titel (wenn abweichend)	Regie	Genre	Mischtonmeister	Studio	Land*	Verleih
3	-	Tom Tykwer	Drama	Matthias Lempert	Elektrofilm, Berlin	D	X Verleih
8 Femmes	8 Frauen	François Ozon	Komödie	-	-	F	Constantin
Alexander	-	Oliver Stone	Action	Paul Massey (US), Vincent Arnardi, Bruno Tarrière et al.	Auditoriums de Boulogne	F	Constantin
Alle Anderen	-	Maren Ade	Drama	Florian Beck, Matthias Schwab	BASIS, Berlin	D	Prokino
Another Year	-	Mike Leigh	Komödie, Drama	Marc Paterson, Nigel Stone, Andy Thompson	Goldcrest, London	UK	Prokino
Basic Instinct 2	-	Michael Caton-Jones	Thriller	Paul Massey (US), Chris Burdon	De Lane Lea, London	UK	Constantin
Biutiful	-	Alejandro González Iñárritu	Drama	Leslie Shatz (US), Bob Beemer (US), Jon Taylor (US)	-	MX	Prokino
Blutbrüder	-	Özgür Yildirim	Musik, Komödie	Richard Borowski	Loft Studios, Hamburg	D	Constantin
Caotica Ana	Chaotisch Ana	Julio Medem	Drama	-	Pecera Estudio, Madrid	E	Prokino
Chéri	-	Stephen Frears	Komödie	Mike Prestwood Smith, Doug Cooper	Real World Studios, Wiltshire	UK	Prokino
Deutschland 09	-	Fatih Akin, Tom Tykwer u.a.	Drama (div.)	Martin Steyer, Matthias Lempert u.a.	RuhrSound Studios, Dortmund	D	Piff! Medien
Die Bluthochzeit	-	Dominique Deruddere	Action, Drama	Martin Steyer	RuhrSound Studios, Dortmund	D	Constantin
Elementarteilchen	-	Oskar Roehler	Drama	Martin Steyer	RuhrSound Studios, Dortmund	D	Constantin
Endlich	-	Katja Dringenberg, Chr. Voss	Dokumentarfilm	Matthias Lempert	Studio Matthias Lempert	D	X Verleih
Erkan & Stefan 2	-	Axel Sand	Komödie	Jürgen Roth	Studio 507, München	D	Constantin
Glück	-	Doris Dörrie	Drama	Tschangis Chahrokh	ARRI Sound, München	D	Constantin
Gomorrha	-	Matteo Garrone	Drama, Action	Leslie Shatz	Wildfire Studios, LA	US	Prokino
Hachi: A Dog's Tale	Hachiko - Eine Wunderbare Freundschaft	Lasse Hallström	Drama	Dave Paterson, Robert Fernandez	Sound One, NY	US	Prokino
Hallam Foe	-	David Mackenzie	Drama	Kahl Henderson, Chris Sinclair	Goldcrest, London	UK	Prokino
Heartbreak Hotel	Schwedisch für Fortgeschrittene	Colin Nutley	Komödie	Gábor Pasztor	-	S	Prokino
Immortals	Krieg der Götter	Tarsem Singh	Action	Leslie Shatz, Chris David, Gabriel J. Serrano	Wildfire Studios, LA	US	Constantin
Kammerflimmern	-	Hendrik Hölzemann	Drama	Abi Schneider	Cinepostproduktion	D	Constantin
La Clinique de l'Amour!	-	Yesim Ustaoglu	Komödie	Bruno Tarrière	Polyson, Paris	F	Match Factory
La Mome	La Vie en Rose	Olivier Dahan	Drama, Musik	Marc Doisne	-	F	Constantin
Le Code a Changé	Affären à la carte	Danièle Thompson	Komödie, Drama	Vincent Arnardi, Julien Perez	-	F	Prokino
Lian ai zhong de Bao Bei	Baober in Love	Shaohong Li	Drama	Bruno Tarrière	Auditoriums de Boulogne	F	Guangdong Wai Jia
Lissi und der wilde Kaiser	-	Michael Herbig	Komödie	Michael Kranz, Benjamin A. Rosenkind	Cinepostproduktion	D	Constantin
Little Thirteen	-	Christian Klandt	Drama	Silvio Naumann	HFF Potsdam-Babelsberg	D	X Verleih
Lost in Translation	Zwischen den Welten	Sofia Coppola	Drama	Robert Knox, Richard Beggs, Kent Sparling	-	US	Constantin
Michael Clayton (DUB)	-	Tony Gilroy	Thriller	Michael Barry, Andrew Kris	-	US	Constantin
Michael Clayton (OV)	-	Tony Gilroy	Thriller	Michael Barry, Andrew Kris	-	US	Constantin
Mine Vaganti	Männer al dente	Ferzan Ozpetek	Komödie, Drama	Roberto Moroni	-	I	Prokino
Ödland	-	Anne Kodura	Dokumentarfilm	Gregor Bonse	ROTOR Studio F, Babelsberg	D	Die Zone
Pandorum	-	Christian Alvart	Action, Horror	Tschangis Chahrokh, Michael Hinreiner	ARRI Sound, München	D	Constantin
Perfume - Story of a Murderer	Das Parfum - Die Geschichte eines Mörders	Tom Tykwer	Drama	Matthias Lempert, Michael Kranz, Benjamin A. Rosenkind	Cinepostproduktion	D	Constantin
Precious	-	Lee Daniels	Drama	Brian Bowles, Robert Hein, Reilly Steele, Robert Fernandez	Sound One, NY	US	Prokino
R U There	-	David Verbeek	Drama, Action	Bruno Tarrière	Cinéphase Audi 3, Paris	F	Filmmus Distributie
Rammbock	-	Marvin Kren	Thriller, Horror	Hanse Warns	Tonstudio Hanse Warns, Berlin	D	Filmgalerie 451
Resident Evil 3: Extinction	-	Russell Mulcahy	Action, Horror	Chris Jenkins (US), Frank A. Montaño (US)	Paramedia Studios, Montreal	Can	Constantin
Step Up 3D	-	Jon M. Chu	Musik, Drama	David E. Fluhr, Dean A. Zupancic	-	US	Constantin
Step Up Revolution	Step Up - Miami Heat	Scott Speer	Musik, Drama	Tony Lambert, Michael Minkler	Todd-AO, CA	US	Constantin
Tamara Drewe	Immer Drama um Tamara	Stephen Frears	Komödie, Drama	Mike Prestwood Smith, Chris Treble	Pepper Post, London	UK	Prokino
The Dark	-	John Fawcett	Thriller, Horror	Howard Bargroff, Dan Johnson,	Videosonics Cinema Sound, London	UK	Constantin
The International	-	Tom Tykwer	Action, Thriller	Matthias Lempert	Elektrofilm, Berlin	D	Sony Pictures
The Passion of the Christ	Die Passion Christi	Mel Gibson	Drama	Kevin O'Connell, Scott Millan, Bob Beemer,	Sony Pictures Studios, LA	US	Constantin
The Three Musketeers (DUB)	Die Drei Musketiere	Paul W.S. Anderson	Action	Michael Kranz, Andrew Stirk (UK), Benjamin A. Rosenkind	-	D	Constantin
The Three Musketeers (OV)	Die Drei Musketiere	Paul W.S. Anderson	Action	Michael Kranz, Andrew Stirk (UK), Benjamin A. Rosenkind	-	D	Constantin
Traumschiff Surprise	-	Michael Herbig	Komödie	Michael Kranz, Benjamin A. Rosenkind	Cinepostproduktion	D	Constantin
Türkisch für Anfänger	-	Bora Dagtekin	Komödie	Hubert Bartholomae	-	D	Constantin
Vorstadtkrokodile	-	Christian Ditter	Kinderfilm	Tschangis Chahrokh	ARRI Sound, München	D	Constantin

(\*bei internationalen Produktionen Land der Mischung)

**Tabelle 2: Filmdatenbank, Allgemeine Informationen**

Originaltitel	Länge	Version / Sprache	Mix	Quelle	Ref.	Int. LEQ(M)	Int. LEQ(A) (dBFS)	Int. 1770-1 (LKFS)	Int. 1770-2 (LUFS)	Dialoglevel LEQ(A) (dBFS)	Dialoglevel 1770-1 (LKFS)	Max.S-T. (LUFS)	Max.M. (LUFS)	Max.TP (dBTP)	LRA (LU)	Sprache (%)
3	2h00m17s	OV dt.	PM	Festplatte		72	-31,2	-30,0	-28,3	-30,6	-29,9	-14,2	-12,1	-0,6	19,6	30,0
8 Femmes	1h50m35s	OV fr.	PM	DA88		80	-23,9	-24,0	-22,4	-23,5	-23,9	-11,7	-4,9	0,7	15,7	59,0
Alexander	2h48m14s	OV fr.	PM	DA88	n/a	84	-18,4	-18,5	-15,9	-22,4	-23,7	-1,7	1,5	1,2	23,2	44,0
Alle Anderen	1h58m55s	OV dt.		DVD-R		72	-31,9	-31,7	-28,1	-29,3	-29,9	-16,3	-11,9	0,6	23,3	33,0
Another Year	2h04m12s	OV engl.		DVD-R		72	-30,6	-31,3	-29,1	-29,5	-30,4	-18,9	-15,5	-5,6	16,9	58,0
Basic Instinct 2	1h52m06s	DUB dt.	PM	DVD-R		74	-28,8	-28,2	-26,2	-29,2	-29,1	-11,2	-8,5	0,5	18,7	46,0
Biutiful	2h22m21s	OV sp.	PM	DVD-R		80	-22,3	-21,5	-18,5	-22,1	-22,8	-8,3	-5,0	0,2	23,1	27,0
Blutzbrüdad	1h31m19s	OV dt.	PM	Festplatte		81	-22,3	-20,2	-18,6	-24,4	-25,0	-9,8	-7,8	-0,7	19,6	61,0
Caotica Ana	1h53m48s	OV sp.	PM	DVD-R		73	-28,2	-27,2	-25,1	-29,9	-29,6	-10,3	-8,4	-0,2	19,4	25,0
Chéri	1h29m50s	OV engl.	PM	DVD-R		72	-30,3	-30,1	-28,5	-30,8	-31,0	-14,8	-13,1	0,1	15,6	46,0
Deutschland 09	2h31m60s	OV dt.	PM	Festplatte		74	-28,8	-28,7	-26,6	-29,5	-29,4	-12,7	-8,4	-0,6	18,9	52,0
Die Bluthochzeit	1h27m22s	OV dt.	PM	DA88		78	-26,5	-25,9	-23,6	-28,2	-28,9	-9,9	-5,9	0,5	20,5	43,0
Elementarteilchen	1h48m44s	OV dt.	PM	DA88		74	-29,8	-28,9	-26,8	-29,7	-29,4	-14,8	-8,6	0,3	20,9	28,0
Endlich	1h17m08s	OV dt.	PM	Festplatte		71	-31,6	-28,9	-27,6	-31,4	-30,3	-16,0	-14,9	-1,8	15,6	35,0
Erkan & Stefan 2	1h20m40s	OV dt.		DA88		84	-22,4	-21,3	-20,2	-24,4	-24,3	-5,3	-1,7	1,6	15,8	47,0
Glück	1h47m29s	OV dt.	PM	DA88		74	-29,2	-28,5	-25,6	-29,8	-30,0	-12,2	-4,1	0,9	23,6	25,0
Gomorrha	2h11m34s	OV it.	PM	DVD-R		78	-25,2	-22,3	-18,6	-25,5	-26,5	-5,7	-4,1	1,4	23,0	39,0
Hachi: A Dog's Tale	1h29m10s	OV engl.	PM	DVD-R		71	-30,7	-30,4	-28,9	-30,3	-31,5	-15,6	-10,5	-0,9	17,7	29,0
Hallam Foe	1h33m14s	OV engl.	PM	DVD-R		77	-27,5	-26,9	-24,2	-27,2	-28,1	-7,9	-4,9	0,8	22,9	28,0
Heartbreak Hotel	1h37m18s	OV schw.		DVD-R		78	-24,7	-24,9	-22,8	-27,1	-28,5	-10,5	-5,0	-0,5	20,8	54,0
Immortals	1h51m12s	OV engl.	PM	Festplatte		88	-15,7	-14,0	-10,7	-19,3	-19,8	1,2	4,5	3,9	24,6	17,0
Kammerflimmern	1h37m04s	OV dt.	PM	DVD-R		68	-34,3	-33,5	-30,8	-35,2	-36,4	-17,4	-12,5	-4,1	23,5	25,0
La Clinique de l'Amour!	1h23m56s	OV trk.	PM	USB Flash		78	-25,0	-23,9	-22,1	-26,4	-26,4	-11,2	-9,6	-2,2	18,7	37,0
La Mome	2h09m52s	OV fr.		DA88		80	-21,6	-22,2	-19,8	-22,6	-23,5	-11,0	-8,9	0,5	21,2	41,0
Le Code a Changé	1h36m05s	OV fr.		DVD-R		74	-27,9	-27,7	-26,2	-30,1	-30,2	-14,0	-11,5	-2,6	17,1	64,0
Lian ai zhong de Bao Bei	1h43m38s	OV chn.(m)	PM	USB Flash		76	-27,5	-26,4	-23,1	-26,4	-29,1	-12,0	-6,9	0,2	23,9	28,0
Lissi und der wilde Kaiser	1h21m25s	OV dt.	PM	DVD-R		86	-17,4	-16,7	-15,7	-20,3	-20,9	-4,8	-2,3	1,1	15,2	65,0
Little Thirteen	1h37m00s	OV dt.	PM	Festplatte		77	-26,3	-24,8	-21,5	-28,7	-29,1	-11,3	-7,8	1,5	24,9	27,0
Lost in Translation	1h38m04s	DUB dt.	PM	DA88		75	-26,2	-26,2	-23,8	-28,1	-29,2	-7,8	-6,4	0,0	22,1	49,0
Michael Clayton (DUB)	1h54m52s	DUB dt.	PM	DVD-R		72	-29,6	-28,0	-26,4	-28,8	-29,0	-14,8	-10,1	0,1	17,9	53,0
Michael Clayton (OV)	1h54m52s	OV engl.	PM	DA88		72	-29,9	-28,1	-26,4	-29,5	-29,8	-14,8	-10,3	-1,2	18,5	45,0
Mine Vaganti	1h48m47s	OV it.		DVD-R		78	-27,9	-25,3	-23,1	-25,4	-27,9	-12,1	-9,9	0,0	21,9	45,0
Ödland	1h19m00s	OV dt.	PM	Festplatte		72	-30,3	-31,4	-29,7	-29,1	-29,9	-19,8	-14,2	0,0	18,0	49,0
Pandorum	1h47m54s	OV engl.	PM	DA88	n/a	86	-18,4	-17,2	-14,8	-24,6	-24,5	-1,6	1,1	3,1	22,0	25,0
Perfume - Story of a Murderer	2h21m28s	DUB dt.	PM	DVD-R		78	-24,2	-24,2	-22,0	-25,3	-25,9	-7,0	-4,1	1,2	21,3	27,0
Precious	1h46m02s	OV	PM	DVD-R		74	-29,1	-29,4	-27,0	-29,1	-29,9	-14,0	-10,8	0,2	19,7	51,0
R U There	1h26m37s	OV holl.	PM	USB Flash		72	-30,7	-30,3	-26,5	-32,9	-33,2	-12,6	-9,3	-1,1	26,5	17,0
Rammbock	1h04m05s	OV dt.	PM	Festplatte		82	-21,6	-20,7	-16,9	-27,6	-28,3	-7,6	-5,9	0,4	25,8	23,0
Resident Evil 3: Extinction	1h34m09s	DUB dt.	PM	DA88		88	-16,8	-16,5	-14,1	-20,6	-21,5	-3,9	0,5	0,3	22,8	33,0
Step Up 3D	1h47m11s	DUB dt.	PM	Festplatte		83	-19,6	-17,5	-15,1	-24,9	-25,4	-8,8	-4,4	0,1	22,4	39,0
Step Up Revolution	1h38m56s	OV engl.	PM	Festplatte		84	-19,3	-18,5	-16,5	-25,1	-25,4	-7,7	-4,8	0,5	21,6	28,0
Tamara Drewe	1h46m49s	OV	PM	DVD-R		77	-26,4	-26,0	-23,6	-28,3	-29,3	-6,8	-2,7	2,7	20,9	56,0
The Dark	1h27m58s	OV engl.	PM	DA88		78	-25,8	-25,1	-21,5	-26,4	-27,9	-8,1	-3,4	0,5	25,1	30,0
The International	1h57m52s	OV engl.	PM	Festplatte		82	-23,5	-23,0	-20,1	-29,3	-29,1	-4,6	-2,4	2,4	21,3	32,0
The Passion of the Christ	2h01m42s	OV hebr.		DA88		81	-23,1	-22,5	-21,1	-24,5	-25,7	-8,5	-5,8	1,2	18,9	29,0
The Three Musketeers (DUB)	1h50m35s	DUB dt.	PM	Festplatte		84	-19,9	-18,5	-16,0	-25,6	-25,3	-5,9	-3,0	1,2	22,8	49,0
The Three Musketeers (OV)	1h50m35s	OV engl.	PM	Festplatte		84	-20,0	-18,5	-16,0	-25,3	-25,6	-5,9	-3,0	1,4	22,8	49,0
Traumschiff Surprise	1h23m47s	OV dt.		DVD-R		82	-22,1	-21,6	-20,6	-24,0	-24,6	-10,0	-4,8	2,7	16,2	59,0
Türkisch für Anfänger	1h48m37s	OV dt.	PM	Festplatte		79	-25,3	-24,2	-22,3	-27,0	-27,6	-8,8	-7,0	1,7	19,8	54,0
Vorstadtkrokodile	1h36m35s	OV dt.	PM	DA88		80	-24,2	-23,5	-21,6	-25,6	-26,6	-10,2	-6,3	0,4	19,9	40,0

Tabelle 3: Filmdatenbank, Messungsergebnisse



## 3.5. Auswertung

### 3.5.1. Objektive Auswertung der Datenbank

Versucht man zunächst, sich eine Übersicht zur aktuellen Kinolautstärke-Situation zu verschaffen, ohne dabei subjektive Hörtests der Filme miteinzubeziehen, sollten als Erstes die  $L_{EQ(M)}$ -Werte begutachtet werden, da sie bereits im Kino-Kontext etabliert sind und am ehesten mit einer Hörerfahrung verknüpft werden können.

Die Spannweite erstreckt sich bei den gemessenen Langfilmen von  $L_{EQ(M)}$  68 bis 88, womit vier der gemessenen Filme sogar durchschnittlich lauter sind als für Trailer zugelassen. Ein solches Ergebnis erscheint z.B. unter Berücksichtigung der Einschätzung von Matthias Lempert, der  $L_{EQ(M)}$  Maximalwert von 85 sei seinerseits mindestens 10 dB zu laut,<sup>3</sup> als äußerst bedenklich. 12 Filme sind andererseits mehr als 10 dB leiser als TASA-konforme Trailer. Das bedeutet, dass solche Filme selbst auf Dolby Level 10 (Kinoprozessor maximal aufgedreht) noch nicht so laut wären wie übliche Trailer. Dabei sind Genres wie Action, Thriller und Horror wie zu erwarten tendenziell lauter, dicht gefolgt von Komödie und Musikfilm, wobei beim Drama die leisesten Filme zu finden sind. Eine nationale Verteilung lässt sich nicht ausmachen. Zwar wurden die zwei lautesten Filme von US-amerikanischen Mischtonmeistern gemischt, jedoch handelt es sich dabei um Produktionen, die eine solche Tendenz auch dramaturgisch begründen. Sonstige US-amerikanische Filme verteilen sich ähnlich wie solche anderer Nationalitäten gleichmäßig über den gesamten  $L_{EQ(M)}$ -Lautheitsbereich.

Als weiterer Anhaltspunkt zur Auswertung der Kennwerte kann die Kontinuität zwischen Mischungen gleicher Mischtonmeister herangezogen werden. Geht man davon aus, dass ein Mischtonmeister bei einer von ihm gewählten Abhörlautstärke - z.B. Dolby Level 7 - arbeitet, und unter der Bedingung, dass er im Beisein von Regisseuren und Produzenten sowie ggf. im Team mit weiteren Mischtonmeistern seine eigenen Präferenzen umsetzen kann bzw. darf, dann sollten diese Mischungen untereinander eine deutliche Lautstärke-Kontinuität aufweisen, die sich auch in den Messergebnissen widerspiegelt.

Inwieweit die oben genannten Bedingungen, d.h. die Möglichkeit nach eigenen Präferenzen zu mischen, bei diesen Filmen erfüllt wurden, bleibt ungewiss. Jedoch zeichnen sich bei den drei Filmen von Martin Steyer erstaunlich durchgängige Dialoglevel (nach BS.1770-1) mit lediglich 0,5 LU Abweichung ab. Ebenfalls weisen

vier Filme von Matthias Lempert eine solche Präzision im Dialoglevel auf: Obwohl Action wie auch Drama und Dokumentarfilm vertreten sind, beträgt die maximale Abweichung 1,1 LU, wobei sich die Werte beider Mischtonmeister bei ca. -29 LKFS nahezu entsprechen („Deutschland 09“ wurde unter Mitwirkung beider Mischtonmeister gemischt). Gleichzeitig schwanken die über die gesamte Programmdauer integrierten Lautheitswerte bei diesen Filmen nach  $L_{EQ(A)}$  und BS.1770-1 mit bis zu 6 dB bzw. LU, nach BS.1770-2 (also EBU R 128) mit bis zu 8 LU und nach  $L_{EQ(M)}$  mit mehr als 10 dB.

„Das Parfum“, ein Film der ebenfalls von Matthias Lempert aber im Team mit Michael Kranz und Benjamin A. Rosenkind gemischt wurde, hat ein höheres Dialoglevel von ca. -26 LKFS, das seinerseits wiederum mit lediglich 0,3 LU Abweichung dem von „The Three Musketeers“ entspricht, der ebenfalls von Michael Kranz und Benjamin A. Rosenkind gemischt wurde. Lauter wurden die Dialoge von diesen Mischtonmeistern mit ca. -21 LKFS in „Lissi und der Wilde Kaiser“, einer deutschen Komödie gemischt. Eine ähnliche Tendenz ist bei Bruno Tarrière festzustellen, dessen französische Komödie „La clinique de l’Amour“ wenn auch nur um 2,7 LU lautere Dialoge aufweist als das von ihm gemischte chinesische Drama „Baober in Love“. Ein weiteres diesmal niederländisches Action-Drama „R U There“ von Tarrière hat wiederum ein deutlich niedrigeres Dialoglevel, worauf zu einem späteren Zeitpunkt noch einmal eingegangen wird. Zwei Filme von Mike Prestwood Smith weisen wieder nahezu identische Dialoglevel auf, während Mischungen anderer Tonmeister - ob sie nun alleine oder in Teams arbeiteten - sich deutlich voneinander unterscheiden.<sup>4</sup> Die Werte für Dialoglevel nach  $L_{EQ(A)}$  verhalten sich ähnlich zu jenen nach BS.1770-1 und fallen durchschnittlich etwa 0,6 dB lauter aus.

Die über den gesamten Film integrierten Programmlautheiten nach  $L_{EQ(M)}$ ,  $L_{EQ(A)}$ , BS.1770-1 und BS.1770-2 weisen bei Mischungen gleicher Mischtonmeister keine Parallelen auf, insbesondere dann nicht, wenn es sich um unterschiedliche Genres handelt. Eine Normalisierung anhand dieser Parameter erscheint also bereits nach dieser Beobachtung nicht sinnvoll, würde sie doch zu Screenings führen, bei denen Filme gleicher Mischtonmeister auf unterschiedlichen Dolby Levels laufen würden.

Wenn über Dialoglevel im Kino gesprochen wird, dann tritt auch häufig die Frage auf, ob und wie sie sich bei Originalversionen und Synchronfassungen unterscheiden. Der Verdacht besteht, dass die Mischungs-Stilistik von synchronisierten Filmen - und speziell zu laut gemischte Dialoge - eine veränderte Erwartungshaltung bei Zuschauern in Ländern provoziert, wo diese Praxis verbreitet

ist. Zwar sind sieben der gemessenen Filme Synchronfassungen, aus dem Vergleich der absoluten Dialoglevel lassen sich aber kaum Schlüsse ziehen, da diese Filmübergreifend zu sehr schwanken und über eine Spanne von 16,6 LU verteilt sind. Zwei Filme liegen jedoch sowohl als Original wie auch synchronisiert vor:

Bei „Michael Clayton“ stimmen Musik und Effekte in ihrer Lautstärke bei beiden Versionen überein, d.h. es wurden für die Synchronfassung keine absoluten Änderungen am International Tape vorgenommen. Auch die Programmlautheiten nach BS.1770-2 sind mit -26,4 LKFS identisch. Laut DMM2 sind die Dialoge in der Synchronfassung durchschnittlich lediglich 0,8 LU lauter. Beide „The Three Musketeers“ Versionen enthalten ebenfalls gleichlaute Musik und Effekte; sie weisen mit -16,0 LKFS gleiche Programmlautheiten auf, wobei die synchronisierten Dialoge hier mit 0,3 LU nur marginal lauter sind. Während Dialogue Intelligence bei „Michael Clayton“ 8% mehr Dialog in der synchronisierten Fassung als in der Originalversion erkannte, sind die erkannten Sprachanteile bei „The Three Musketeers“ identisch. Der Verdacht eines deutlich gesteigerten Dialoglevels bei synchronisierten Filmen lässt sich also anhand der verfügbaren Daten letztlich weder bestätigen noch ausschließen.

Bei den zeitabhängigen Parametern wurden die größten Maximalwerte mit +1,2 LUFS Short-Term Loudness, +4,5 LUFS Momentary Loudness sowie +2,4 dBTP True Peak Level beim Film „Immortals“ gemessen. Während besonders hohe Short-Term und Momentary Lautheits-Werte Genres wie Action und Horror zugeordnet werden können, nutzen fast alle Filme den gesamten technischen Aussteuerungs-Bereich aus; knapp 70% der Filme überschreiten sogar 0 dBTP True Peak. Die Loudness Range gibt z.T. ebenfalls vage Hinweise auf Genres, z.B. weist eine kleine Loudness Range oft auf eine Komödie hin, während sowohl Dramen als auch Actionfilme große Werte haben können. Auch der von Dialogue Intelligence gemessene Sprachanteil lässt sich gewissermaßen mit Stilistik und Genre des Films verknüpfen: So spricht den Ergebnissen zufolge eine hohe Sprachdichte deutlich für Komödie und eine niedrigere deutlich mehr für Drama.

Abbildung 11 verdeutlicht, wie die Programmlautheit bei Genres wie Dokumentarfilm und Drama dem Dialoglevel am ehesten entspricht oder dieses sogar unterschreitet, während die Differenz der beiden Werte bei Komödien zunimmt und bei Action sowie Musikfilm am größten ist. Eine Normalisierung nach Programmlautheit hätte also gerade bei letzteren Genres verhältnismäßig leise Dialoge zur Folge.

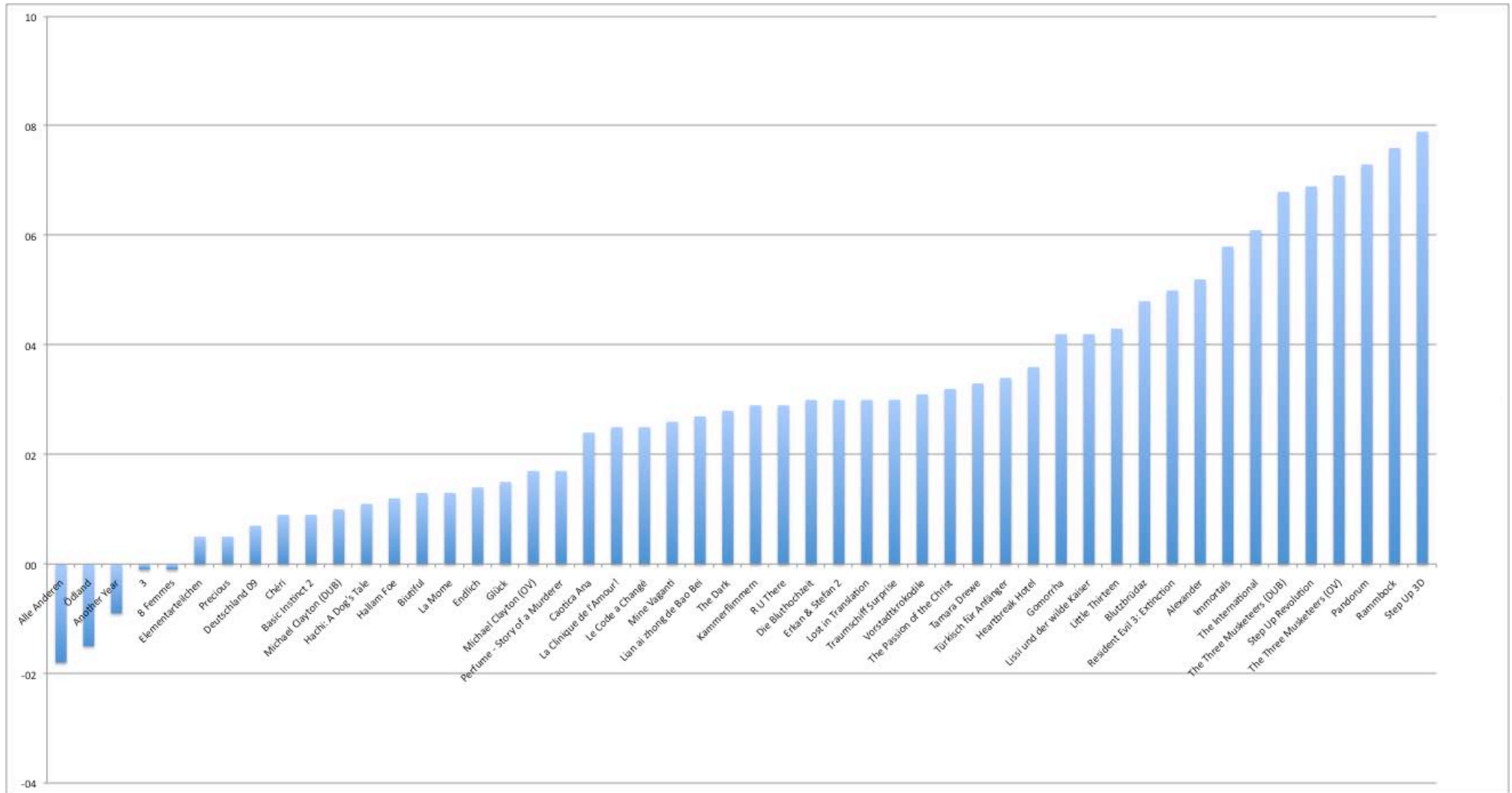


Abbildung 11: Differenz von Programmlautheit und Dialoglevel nach BS.1770-1 (LU)

### 3.5.2. Auswertung der Graphen

An den Graphen des Dolby Media Meter lässt sich besonders gut erkennen, wie die Differenz zwischen integriertem Dialoglevel und Programmlautheit zustande kommt.

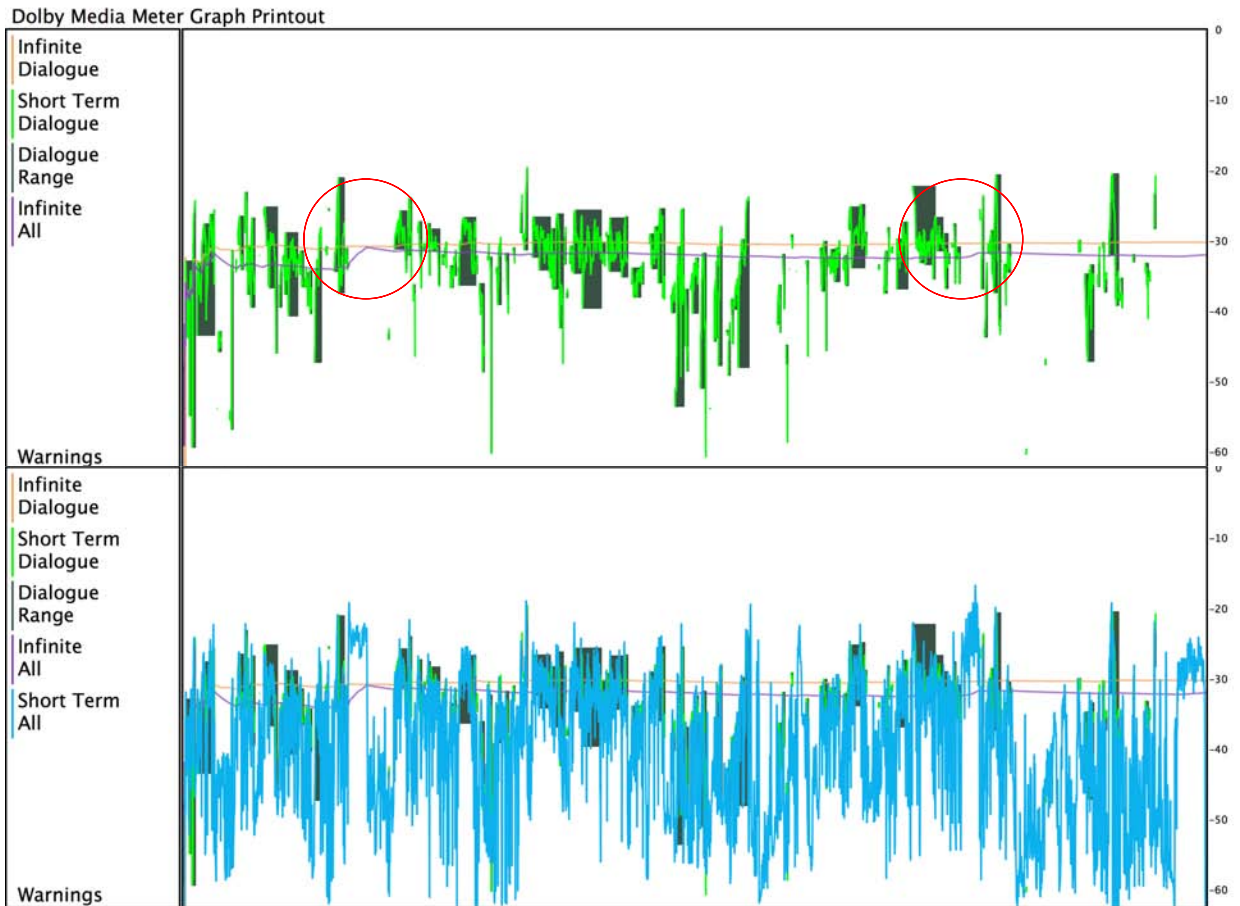


Abbildung 12: Entwicklung der Lautheit nach BS.1770-1 im Film „Alle Anderen“

Abbildung 12 zeigt die im Film „Alle Anderen“ gemessenen Dialogpassagen (Short-Term Dialogue, grün) und die Entwicklung der ungegateden Programmlautheit (Infinite All, violett) sowie des über den gesamten Film integrierten Dialoglevels (Infinite Dialogue, orange). Der untere Teil der Grafik zeigt zusätzlich die Short-Term Loudness (Short-Term All, blau). Die Dialoge weisen eine typische Kinodynamik auf und nach einer kurzen Einschwingphase pendelt sich das Dialoglevel bei ca. -30 LKFS ein. Die Programmlautheit entwickelt sich etwas unregelmäßiger: Nach einer anfänglichen Einschwingung treibt im ersten Viertel des Films eine Musiksequenz (erkennbar durch Abwesenheit von Dialog sowie gesteigerte und kontinuierliche Short-Term Loudness) die Programmlautheit in die Höhe. Sie sinkt anschließend durch leise Passagen zwischen den Dialogen wieder ab und steigt im letzten Viertel durch eine weitere Musiksequenz noch einmal kurz an.

Letztlich endet die Programmlautheit auf einem Wert unterhalb des Dialoglevels, das seinerseits durch die Musikpassagen nicht manipuliert wurde.

Beim Film „The International“ zeigt sich ein für Action-Filme typischer Lautstärke-Verlauf. Trotz dynamischer Sprachpassagen steht das von Dialogue Intelligence mit BS.1770-1 erfasste Dialoglevel nahezu statisch bei ca. -29 LKFS. Die Programmlautheit hingegen, schlägt mehrmals aus, z.B. bei der Schießerei im Guggenheim Museum im dritten Viertel des Films. Obwohl die Auswirkung lauter Passagen auf die integrierte Lautheit statistisch gesehen zum Ende hin abnehmen muss, da bereits über einen größeren Zeitraum gemittelt wurde, ist der Effekt deutlich im Graph zu erkennen. Am Ende liegt die ungegatede Programmlautheit 6 LU über dem Dialoglevel (Abbildung 13), die gatede sogar 9 LU (s. Datenbank). Es wird also deutlich, dass einzelne Rollen oder für sich gemessene Passagen dieses Filmes zwar identische Dialoglevel, aber sehr unterschiedliche Programmlautheits-Werte zur Folge hätten.

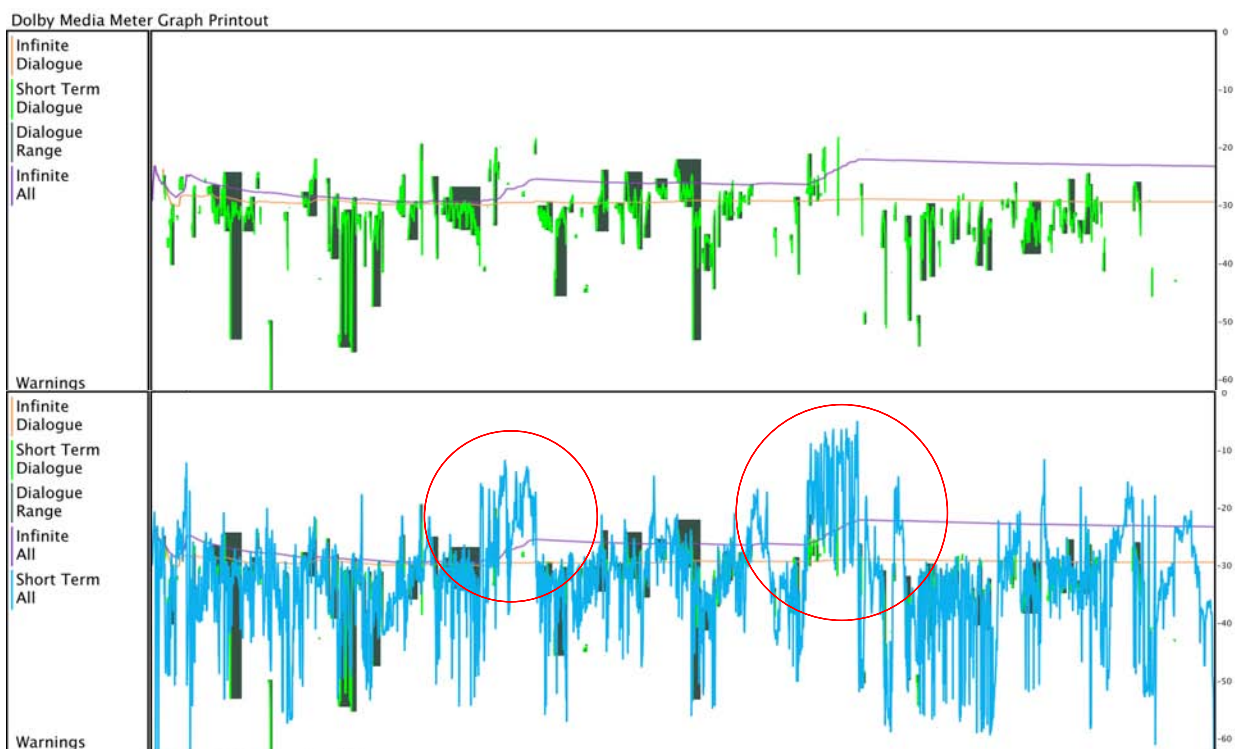


Abbildung 13: Entwicklung der Lautheit nach BS.1770-1 im Film „The International“

Bei der Komödie „Lissi und der Wilde Kaiser“ zeigen sich ein besonders hoher Sprachanteil und erstaunlich kompakt gemischte Dialoge (Abbildung 14).

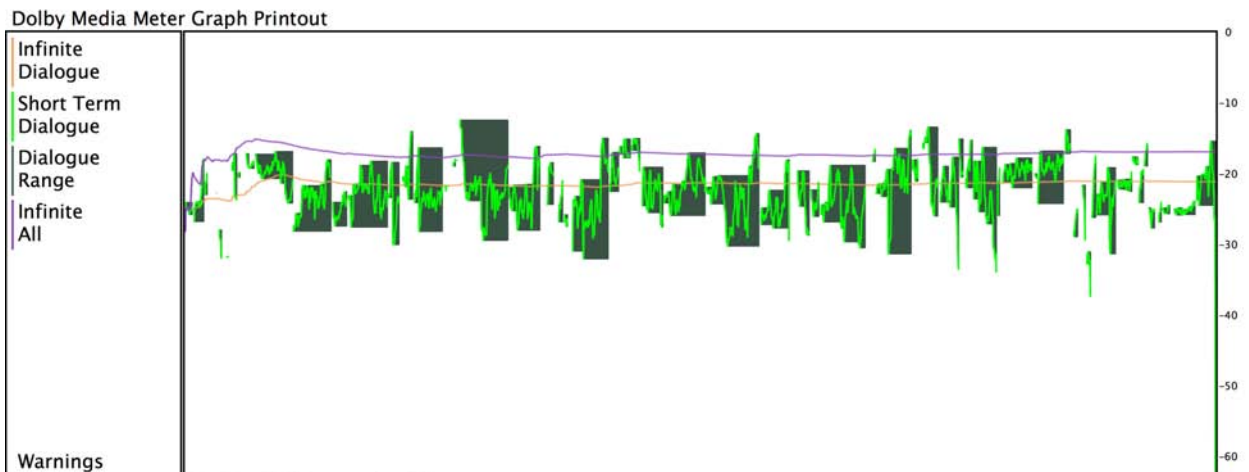


Abbildung 14: Entwicklung der Lautheit nach BS.1770-1 im Film „Lissi und der wilde Kaiser“

Der vom NUGEN LMB erzeugte Graph zum True Peak Level (blau, Abbildung 15) lässt zudem darauf schließen, dass hier massive Dynamik-Einengung betrieben wurde. Die Wellenform der Mischung gleicht eher der einer DVD-Fassung, wenngleich die Kennzeichnung „SRD“ auf den Originaldaten das ausschließt. Der geringere zur Verfügung stehende Headroom, wie er beim Anheben des Dialoglevels entsteht (hier ca. -21 LKFS), wird in beiden Graphen deutlich.



Abbildung 15: Short-Term Loudness und True Peak Level im Film „Lissi und der wilde Kaiser“

### 3.5.3. Beurteilung der Dialogue Intelligence Performance

Am Film „Blutzbrüdad“ erkennt man gut wie die Dolby Spracherkennung nicht nur zeitlich zwischen Sprache und Nicht-Sprache gattet, sondern auch innerhalb der 5.1-Kanäle unterscheidet. Im zweiten sowie im vierten Viertel des Films gibt es Passagen, dessen Short-Term Loudness deutlich hervorsteht. Gleichzeitig wird hier Dialog erkannt, wodurch das integrierte Dialoglevel (Infinite Dialogue) eigentlich



beeinflusst werden müsste. Der erkannte Dialog liegt jedoch deutlich unter der gemessenen Short-Term Loudness dieser Stellen und bewirkt somit nur eine minimale Veränderung des integrierten Dialoglevels. Es handelt sich hierbei um Rap-Musik. Für die Berechnung der Short-Term Loudness werden bekanntlich alle Kanäle außer der LFE gemessen. Dialogue Intelligence wertet aber nur den Center Kanal aus, da nur in diesem Dialog erkannt wurde, wodurch wesentlich kleinere Lautheits-Werte zustande kommen, da die meiste Energie der Musik in den Kanälen Links und Rechts liegt.

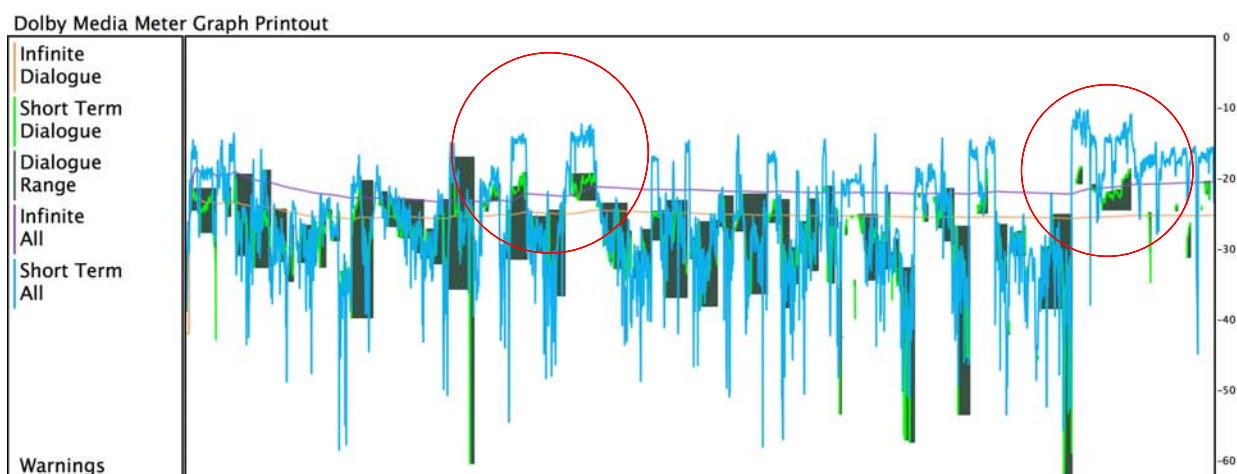


Abbildung 16: Spracherkennung von Rap-Passagen im Film „Blutzbrüdz“

Eigentlich sollten Rap oder sonstige rhythmisierte Sprache bzw. Gesang gar nicht mit in die Dialog-Messung einfließen, was durch die Spracherkennungs-Features Nr. 6 und 7 erreicht werden soll (Kapitel 2.3.2.). Diese Differenzierung funktioniert hier nicht; dennoch ist die Performance von Dialogue Intelligence an den genannten Stellen als insgesamt positiv zu werten, da eine Verfälschung des Dialoglevels durch Inter-Channel-Gating verhindert wird. Wenngleich Dolby keine Angaben dazu macht, wie der Dialogue Intelligence Algorithmus im Einzelnen beim DMM2 implementiert wurde, wird das oben beschriebene Inter-Channel-Gating durch eine schematische Grafik belegt, die Dolby zur lizenzierten Benutzung des Algorithmus in Software von Drittanbietern empfiehlt (Abbildung 17).<sup>5</sup> Demnach soll die Lautheit des 5.0-Signals wie nach ITU-R BS.1770-2 gemessen werden. Parallel dazu werden die drei Front-Kanäle durch unabhängige DI-Instanzen geführt, die jeweils einen Verstärker auf 1 (Sprache) oder 0 (keine Sprache) schalten. Die drei Verstärker gaten damit die um ca. zwei Sekunden verzögerten LCR-Signale. Anschließend werden diese summiert und logarithmisiert bevor sie ein Speech-Gate durchlaufen, das unter -70 LKFS alles (also auch Dialog-freie Passagen) gated und welches auch den Sprachanteil berechnet. Dieser kann anschließend dazu benutzt werden, dass z.B.



unter einem bestimmten Sprachanteil-Schwellenwert statt des Dialoglevels die Lautheit nach BS.1770-2 ausgegeben wird. Diese Funktion wird vom DMM2 jedoch nicht unterstützt.

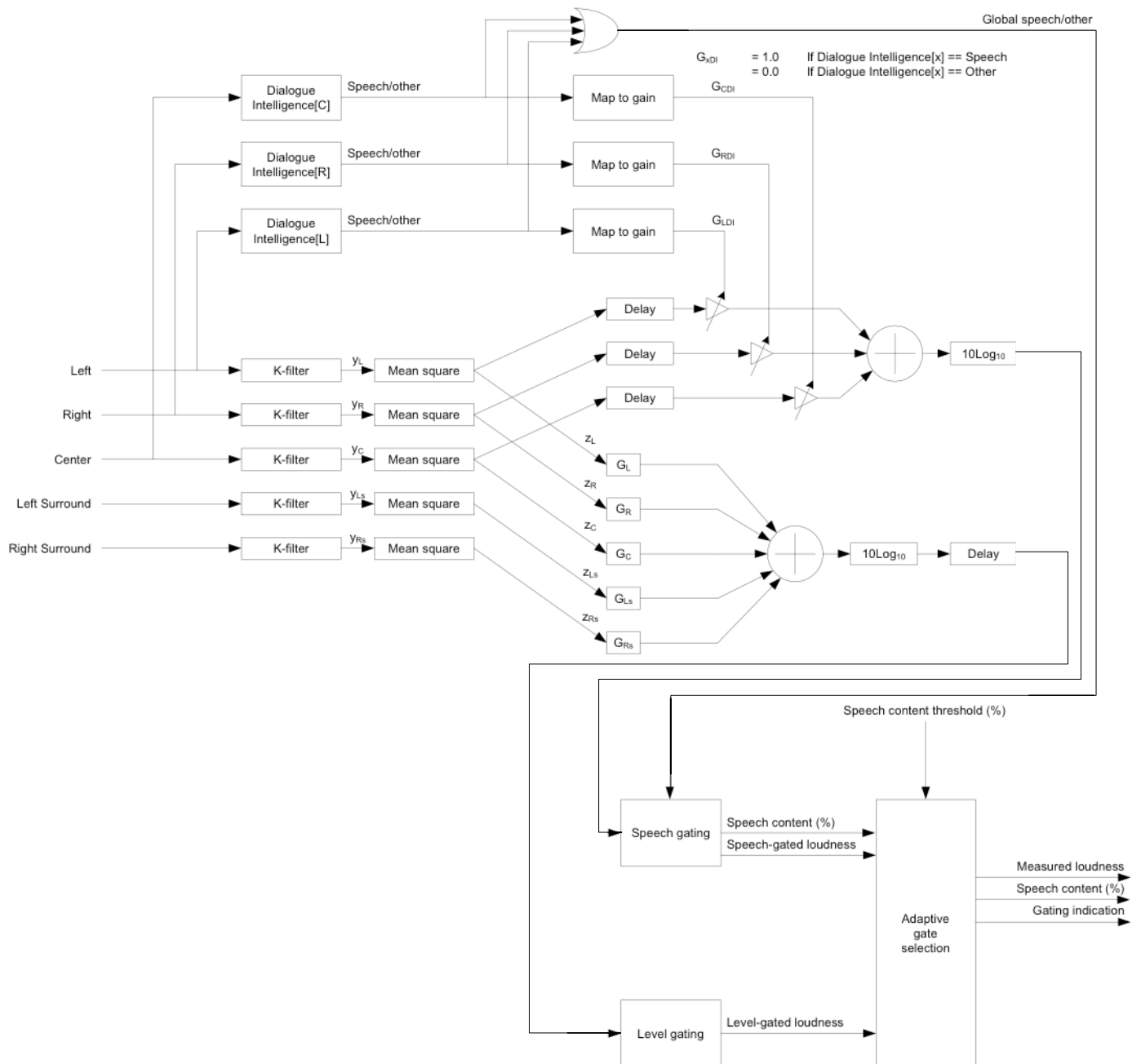


Abbildung 17: Implementierung Dolby Dialogue Intelligence

Als problematisch ist zu erachten, dass das Sprachgate durch seine Funktionsweise z.T. verzögert reagiert, wenn eine Sprachpassage beginnt oder endet. Wechselt das Programm an diesen Stellen von einer sehr leisen Passage zu Dialog, bzw. von Dialog zu leiser Passage (z.B. nur Hintergrund-Atmo), dann fällt dieser Sprung noch in die Dialogwertung, was manche Graphen auf den ersten Blick eher unseriös erscheinen lässt (siehe Abbildung 19, Kapitel 4.1.). Dialogpassagen schwanken dann dem Anschein nach in ihrer Lautheit viel stärker als subjektiv wahrgenommen. Aus diesem Grund wird ein an sich möglicherweise hilfreicher DI-Parameter „Dialogue Range“ völlig unbrauchbar. Dieser Wert gibt die jeweils maximal und minimal gemessene Dialog Short-Term Loudness eines Programms an, was durch den beschriebenen Fehler Ausmaße wie -12 und -68 LKFS annehmen kann. Eine Messung der Dialogue Range wurde daher vollständig aus dieser Studie ausgeklammert. Eine Auswertung des 10. und 95. Perzentils nach Vorbild der Loudness Range wäre für diesen Parameter zweckdienlich.

Im Übrigen macht Dialogue Intelligence gelegentlich Fehldetektionen von Sprache oder Nicht-Sprache in geringem, selten aber auch relevantem Ausmaß (Kapitel 4.1.).

#### ***3.5.4. Vergleich von Messergebnissen und subjektiver Einschätzung***

Wenn man nun die verschiedenen Algorithmen auf ihre Aussagekraft und ihr Potenzial hin untersuchen möchte, eine Lautheits-Normalisierung im Kino zu bewerkstelligen, dann müssen die Werte mit subjektiven Wiedergabe-Erwartungen verglichen werden. Wie bereits erläutert wurden diese Wiedergabe-Erwartungen durch eine subjektive Pegelung des Autors angenähert (Abbildung 18). Werden die gemessenen Programmlautheits-Werte und Dialoglevel mit dieser subjektiven Pegelung verstärkt, dann erhält man zunächst die daraus theoretisch resultierenden Programmlautheiten und Dialoglevel (Tabelle 4). Schwanken diese Werte dann noch stark, bedeutet das, der Algorithmus hat wenig Aussagekraft bzw. er eignet sich schlecht und würde wenig den Vorstellungen eines Zuschauers entsprechen. Die Werte werden aus diesem Grund als Abweichung zu einem Mittelwert dargestellt und eine größtmögliche Abweichung sowie ein Schwankungsbereich wurde errechnet (Tabelle 5).

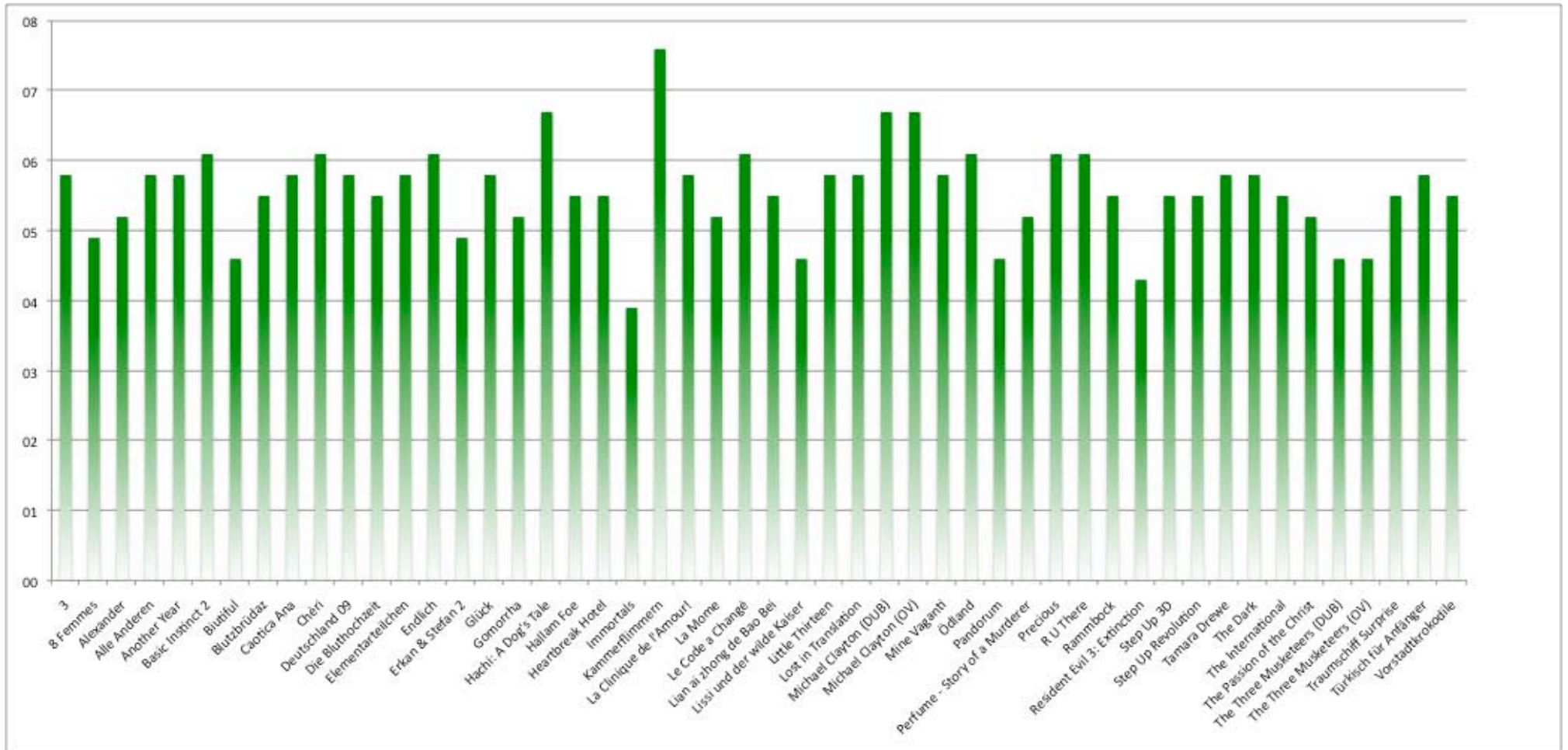


Abbildung 18: Subjektive Pegelung Autor (Dolby Skala)

Originaltitel	Subj. Pegelung Autor (dB rel. Zu Dolby 7)	Result. Int. LEQ(M)	Result. Int. LEQ(A) (dBFS)	Result. Int. 1770-1 (LKFS)	Result. Int. 1770-2 (LUFS)	Result. Dialoglevel LEQ(A) (dBFS)	Result. Dialoglevel 1770-1 (LKFS)
3	-4,0	71	-35,2	-34,0	-32,3	-34,6	-33,9
8 Femmes	-7,0	73	-30,9	-31,0	-29,4	-30,5	-30,9
Alexander	-6,0	78	-24,4	-24,5	-21,9	-28,4	-29,7
Alle Anderen	-4,0	68	-35,9	-35,7	-32,1	-33,3	-33,9
Another Year	-4,0	68	-34,6	-35,3	-33,1	-33,5	-34,4
Basic Instinct 2	-3,0	71	-31,8	-31,2	-29,2	-32,2	-32,1
Biutiful	-8,0	72	-30,3	-29,5	-26,5	-30,1	-30,8
Blutzbrüdad	-5,0	76	-27,3	-25,2	-23,6	-29,4	-30,0
Caotica Ana	-4,0	69	-32,2	-31,2	-29,1	-33,9	-33,6
Chéri	-3,0	69	-33,3	-33,1	-31,5	-33,8	-34,0
Deutschland 09	-4,0	70	-32,8	-32,7	-30,6	-33,5	-33,4
Die Bluthochzeit	-5,0	73	-31,5	-30,9	-28,6	-33,2	-33,9
Elementarteilchen	-4,0	70	-33,8	-32,9	-30,8	-33,7	-33,4
Endlich	-3,0	68	-34,6	-31,9	-30,6	-34,4	-33,3
Erkan & Stefan 2	-7,0	77	-29,4	-28,3	-27,2	-31,4	-31,3
Glück	-4,0	70	-33,2	-32,5	-29,6	-33,8	-34,0
Gomorra	-6,0	72	-31,2	-28,3	-24,6	-31,5	-32,5
Hachi: A Dog's Tale	-1,0	70	-31,7	-31,4	-29,9	-31,3	-32,5
Hallam Foe	-5,0	72	-32,5	-31,9	-29,2	-32,2	-33,1
Heartbreak Hotel	-5,0	73	-29,7	-29,9	-27,8	-32,1	-33,5
Immortals	-11,0	77	-26,7	-25,0	-21,7	-30,3	-30,8
Kammerflimmern	2,0	70	-32,3	-31,5	-28,8	-33,2	-34,4
La Clinique de l'Amour!	-4,0	74	-29,0	-27,9	-26,1	-30,4	-30,4
La Mome	-6,0	74	-27,6	-28,2	-25,8	-28,6	-29,5
Le Code a Changé	-3,0	71	-30,9	-30,7	-29,2	-33,1	-33,2
Lian ai zhong de Bao Bei	-5,0	71	-32,5	-31,4	-28,1	-31,4	-34,1
Lissi und der wilde Kaiser	-8,0	78	-25,4	-24,7	-23,7	-28,3	-28,9
Little Thirteen	-4,0	73	-30,3	-28,8	-25,5	-32,7	-33,1
Lost in Translation	-4,0	71	-30,2	-30,2	-27,8	-32,1	-33,2
Michael Clayton (DUB)	-1,0	71	-30,6	-29,0	-27,4	-29,8	-30,0
Michael Clayton (OV)	-1,0	71	-30,9	-29,1	-27,4	-30,5	-30,8
Mine Vaganti	-4,0	74	-31,9	-29,3	-27,1	-29,4	-31,9
Ödland	-3,0	69	-33,3	-34,4	-32,7	-32,1	-32,9
Pandorum	-8,0	78	-26,4	-25,2	-22,8	-32,6	-32,5
Perfume - Story of a Murderer	-6,0	72	-30,2	-30,2	-28,0	-31,3	-31,9
Precious	-3,0	71	-32,1	-32,4	-30,0	-32,1	-32,9
R U There	-3,0	69	-33,7	-33,3	-29,5	-35,9	-36,2
Rammbock	-5,0	77	-26,6	-25,7	-21,9	-32,6	-33,3
Resident Evil 3: Extinction	-9,0	79	-25,8	-25,5	-23,1	-29,6	-30,5
Step Up 3D	-5,0	78	-24,6	-22,5	-20,1	-29,9	-30,4
Step Up Revolution	-5,0	79	-24,3	-23,5	-21,5	-30,1	-30,4
Tamara Drewe	-4,0	73	-30,4	-30,0	-27,6	-32,3	-33,3
The Dark	-4,0	74	-29,8	-29,1	-25,5	-30,4	-31,9
The International	-5,0	77	-28,5	-28,0	-25,1	-34,3	-34,1
The Passion of the Christ	-6,0	75	-29,1	-28,5	-27,1	-30,5	-31,7
The Three Musketeers (DUB)	-8,0	76	-27,9	-26,5	-24,0	-33,6	-33,3
The Three Musketeers (OV)	-8,0	76	-28,0	-26,5	-24,0	-33,3	-33,6
Traumschiff Surprise	-5,0	77	-27,1	-26,6	-25,6	-29,0	-29,6
Türkisch für Anfänger	-4,0	75	-29,3	-28,2	-26,3	-31,0	-31,6
Vorstadtkrokodile	-5,0	75	-29,2	-28,5	-26,6	-30,6	-31,6
	<b>Mittelwert</b>	<b>73,1</b>	<b>-30,2</b>	<b>-29,4</b>	<b>-27,2</b>	<b>-31,8</b>	<b>-32,3</b>

**Tabelle 4: Nach subj. Pegelung resultierende Programmlautheiten und Dialoglevel mit Mittelwert**

Der größte Schwankungsbereich existiert nach dieser Untersuchung bei der ungegateden Programmlautheit nach BS.1770-1 mit 13,2 LU dicht gefolgt von gatedeter Programmlautheit nach BS-1770-2 mit 13,0 LU. Das bedeutet, dass bei einer „R 128 im Kino“ - also einer Normalisierung nach dem Algorithmus BS.1770-2 - Wiedergabe-Lautstärken entstehen können, die möglicherweise mit mehr als 6 LU sowohl nach oben als auch nach unten von den relativen Wiedergabe-Erwartungen

Originaltitel	Normalisierung nach Int. LEQ(M) (dB)	Normalisierung nach Int. LEQ(A) (dB)	Normalisierung nach 1770-1 (LU)	Normalisierung nach 1770-2 (LU)	Normalisierung nach Dialoglevel LEQ(A) (dB)	Normalisierung nach Dialoglevel 1770-1 (LU)
3	2,1	5,0	4,6	5,1	2,8	1,6
8 Femmes	0,1	0,7	1,6	2,2	-1,3	-1,4
Alexander	-4,9	-5,8	-4,9	-5,3	-3,4	-2,6
Alle Anderen	5,1	5,7	6,3	4,9	1,5	1,6
Another Year	5,1	4,4	5,9	5,9	1,7	2,1
Basic Instinct 2	2,1	1,6	1,8	2,0	0,4	-0,2
Biutiful	1,1	0,1	0,1	-0,7	-1,7	-1,5
Blutbrüder	-2,9	-2,9	-4,2	-3,6	-2,4	-2,3
Caotica Ana	4,1	2,0	1,8	1,9	2,1	1,3
Chéri	4,1	3,1	3,7	4,3	2,0	1,7
Deutschland 09	3,1	2,6	3,3	3,4	1,7	1,1
Die Bluthochzeit	0,1	1,3	1,5	1,4	1,4	1,6
Elementarteilchen	3,1	3,6	3,5	3,6	1,9	1,1
Endlich	5,1	4,4	2,5	3,4	2,6	1,0
Erkan & Stefan 2	-3,9	-0,8	-1,1	0,0	-0,4	-1,0
Glück	3,1	3,0	3,1	2,4	2,0	1,7
Gomorra	1,1	1,0	-1,1	-2,6	-0,3	0,2
Hachi: A Dog's Tale	3,1	1,5	2,0	2,7	-0,5	0,2
Hallam Foe	1,1	2,3	2,5	2,0	0,4	0,8
Heartbreak Hotel	0,1	-0,5	0,5	0,6	0,3	1,2
Immortals	-3,9	-3,5	-4,4	-5,5	-1,5	-1,5
Kammerflimmern	3,1	2,1	2,1	1,6	1,4	2,1
La Clinique de l'Amour!	-0,9	-1,2	-1,5	-1,1	-1,4	-1,9
La Môme	-0,9	-2,6	-1,2	-1,4	-3,2	-2,8
Le Code a Changé	2,1	0,7	1,3	2,0	1,3	0,9
Lian ai zhong de Bao Bei	2,1	2,3	2,0	0,9	-0,4	1,8
Lissi und der wilde Kaiser	-4,9	-4,8	-4,7	-3,5	-3,5	-3,4
Little Thirteen	0,1	0,1	-0,6	-1,7	0,9	0,8
Lost in Translation	2,1	0,0	0,8	0,6	0,3	0,9
Michael Clayton (DUB)	2,1	0,4	-0,4	0,2	-2,0	-2,3
Michael Clayton (OV)	2,1	0,7	-0,3	0,2	-1,3	-1,5
Mine Vaganti	-0,9	1,7	-0,1	-0,1	-2,4	-0,4
Ödland	4,1	3,1	5,0	5,5	0,3	0,6
Pandorum	-4,9	-3,8	-4,2	-4,4	0,8	0,2
Perfume - Story of a Murderer	1,1	0,0	0,8	0,8	-0,5	-0,4
Precious	2,1	1,9	3,0	2,8	0,3	0,6
R U There	4,1	3,5	3,9	2,3	4,1	3,9
Rammbock	-3,9	-3,6	-3,7	-5,3	0,8	1,0
Resident Evil 3: Extinction	-5,9	-4,4	-3,9	-4,1	-2,2	-1,8
Step Up 3D	-4,9	-5,6	-6,9	-7,1	-1,9	-1,9
Step Up Revolution	-5,9	-5,9	-5,9	-5,7	-1,7	-1,9
Tamara Drewe	0,1	0,2	0,6	0,4	0,5	1,0
The Dark	-0,9	-0,4	-0,3	-1,7	-1,4	-0,4
The International	-3,9	-1,7	-1,4	-2,1	2,5	1,8
The Passion of the Christ	-1,9	-1,1	-0,9	-0,1	-1,3	-0,6
The Three Musketeers (DUB)	-2,9	-2,3	-2,9	-3,2	1,8	1,0
The Three Musketeers (OV)	-2,9	-2,2	-2,9	-3,2	1,5	1,3
Traumschiff Surprise	-3,9	-3,1	-2,8	-1,6	-2,8	-2,7
Türkisch für Anfänger	-1,9	-0,9	-1,2	-0,9	-0,8	-0,7
Vorstadtkrokodile	-1,9	-1,0	-0,9	-0,6	-1,2	-0,7
Größte positive Abweichung	5,1	5,7	6,3	5,9	4,1	3,9
Größte negative Abweichung	-5,9	-5,9	-6,9	-7,1	-3,5	-3,4
<b>Schwankungsbereich</b>	<b>11,0</b>	<b>11,6</b>	<b>13,2</b>	<b>13,0</b>	<b>7,6</b>	<b>7,3</b>

**Tabelle 5: Abweichungen vom Mittelwert und resultierender Schwankungsbereich**

eines Zuschauers abweichen würden, vorausgesetzt dieser Zuschauer strebt eine Film-übergreifend kontinuierliche Kino-Lautstärke an. Anders ausgedrückt wären zwischen den Filmen Lautstärke-Sprünge von bis zu 13,0 LU möglich, was sich äquivalent zu 13,0 dB verhält und je nach Definition mit etwa „drei bis vier Mal so laut“ umschrieben werden kann.<sup>6</sup> Jedoch muss bedacht werden was hier genau springt. Ebenso gut könnte ja gesagt werden, dass die Lautstärke kontinuierlich

bleibt, da ja nach BS.1770, einer für Sprache und Musik erprobten Methode normalisiert wird und sich laute und leise Stellen insgesamt wieder ausgleichen. Das Problem liegt darin, dass es für Filme nicht eine einzige allgemeingültige Lautstärke geben kann, da sie dynamisch sind und sich zudem genre-bedingt in ihrer Dynamik unterscheiden. Wenn ein Zuschauer also eine passende Wiedergabe-Lautstärke wählt, ist diese bereits ein Kompromiss, eine Mitte. Bei dieser Untersuchung zeichnet sich die Tendenz ab, dass die Mittelung zwischen lauten und leisen Stellen wie nach BS.1770 für eine optimale Wiedergabe-Lautstärke im Kino wenig geeignet ist, da sie eine „fiktive Mitte“ ist. Sie verschiebt sich in unlogischer Weise durch die Quantität lauter Passagen, selbst wenn diese in der Dramaturgie oder im Genre begründet sind. Dadurch entspricht sie nicht der Idee eines liberalen Gebrauchs von Dynamik, d.h. dem dramaturgischen Einsatz von „Laut“ und „Leise“ im Kino.

Entgegen möglicher Erwartungen fällt der Schwankungsbereich bei Lautheits-Normalisierung nach  $L_{EQ(A)}$  mit 11,6 dB und nach  $L_{EQ(M)}$  mit 11,0 dB kleiner aus. Durchschnittlich resultieren nach der subjektiven Pegelung  $L_{EQ(M)}$ -Werte von ca. 73 dB<sub>(M)</sub>. Das heißt, bei diesem Wert, der ganze 12 dB geringer ist, als die TASA Vorschrift für Trailer, ist laut den subjektiven Einschätzungen des Autors eine angenehme Wiedergabe für auf Dolby Level 7 gezeigte Langfilme zu erwarten. Geht man davon aus, dass Trailer allein aufgrund ihrer dramaturgischen Dichte von lauten Elementen etwas größere Messwerte erreichen müssen als Langfilme, dann erscheint die Einschätzung von Matthias Lempert durchaus realistisch, die TASA Vorschrift sei um bis zu 10 dB zu hoch.<sup>7</sup>


Wenn man hingegen die Dialoglevel der Filme normalisiert, ist der Schwankungsbereich gemessen nach BS.1770-1 mit 7,3 LU am geringsten (gemessen nach  $L_{EQ(A)}$  ist die Schwankung mit 7,6 LU unwesentlich höher). Die größte negative Abweichung, also Filme die nach der Normalisierung zu leise wiedergegeben würden, liegt bei -3,4 LU, die größte positive bei +3,9 LU. Vergleicht man diese Ergebnisse mit der Komfort-Zone nach Riedmiller et al.<sup>8</sup>, dann liegt die untere Abweichung von -3,4 LU noch über dem Wert „leiser, aber akzeptabel“ von -5,4 dB. Die obere Abweichung von +3,9 LU liegt 1,5 LU über dem Wert „lauter, aber akzeptabel“ von +2,4 dB. Damit liegen 96% der gemessenen Filme innerhalb der Komfort-Zone eines Zuschauers, wenn die Dialoglevel auf seine persönliche Ideallautstärke normalisiert werden. Keiner der Filme würde eine Beschwerde des Zuschauers oder ein Verändern der Lautstärke per Fernbedienung bewirken.<sup>9</sup> Geht man davon aus, dass Zuschauer bei typischem TV-Programm selbst unter idealen


Bedingungen keine Lautstärke-Unterschiede unter 1,24 dB ausmachen können,<sup>10</sup> dann sind knapp 50% der Filme nicht merkbar lauter oder leiser als der Durchschnitt. Diese Methode funktioniert also dadurch, dass sie nicht nach einer fiktiven sondern einer realen Mitte, einer der Quasi-Konstanten des Filmtons normalisiert, dem Dialog.

Selbstverständlich schließt das noch nicht aus, dass Toleranzen beim Zuschauer überschritten werden, wenn einzelne Passagen wie Actionszenen im Bezug zum Dialoglevel zu laut gemischt wurden.


---

<sup>1</sup> Am Kinoprozessor entsprechen 0,5 Schritte von 0-4 je 10dB und von 4-10 je 3/5dB


<sup>2</sup> vgl.  Riedmiller, Lyman, Robinson (Dolby Laboratories): Intelligent Program Loudness Measurement and Control - What Satisfies Listeners? S. 5f


<sup>3</sup> vgl.  DUC: Dolby 7. Is everybody mixing at this level? (matt-o-, 14.05.11, 06:05)

<sup>4</sup> nochmals sei an dieser Stelle darauf hingewiesen, dass im Hinblick auf absolute Werte wie Dialoglevel Fehler aufgrund falscher vorliegender Audio-Daten z.B. DVD-Mischungen nicht endgültig ausgeschlossen werden können

<sup>5</sup> vgl.  Dolby Laboratories: Dolby Dialogue Intelligence Reference Code User's Guide

<sup>6</sup> Eine subjektiv empfundene Verdopplung der Lautstärke wird je nach Definition bei 6dB, gelegentlich auch bei 10dB angenommen. Generell sollten diese Einordnungen lediglich als ungefähre Verständnishilfe dienen.

<sup>7</sup> vgl.  DUC: Dolby 7. Is everybody mixing at this level? (matt-o-, 14.05.11, 06:05)

<sup>8</sup> vgl.  Riedmiller, Lyman, Robinson (Dolby Laboratories): Intelligent Program Loudness Measurement and Control - What Satisfies Listeners? S. 5f

<sup>9</sup> Ein leiser bzw. lauter Schalten ist bei TV-Zuschauern außerhalb von +5,6 und -10,2 dB (bezogen auf ihre Ideallautstärke) zu erwarten

<sup>10</sup> Soulodre, Lavoie, Norcross: The Subjective Loudness of Typical Program Material, S. 3

## 4. Konzeption einer Modellnorm

Im vorigen Kapitel wurde die Effektivität verschiedener Lautheits-Normalisierungen im Bezug zur subjektiven Wahrnehmung im Kino getestet. Dabei hat sich die Normalisierung des Dialoglevels, gemessen durch Dolby Dialogue Intelligence mit dem ungegateden Lautheits-Algorithmus BS.1770-1, mit einem Schwankungsbereich von 7,3 LU als beste Methode herausgestellt. Im Folgenden wird untersucht, welche Verbesserungen für eine Modellnorm nötig sind, wie diese implementiert werden können und welche Auswirkungen zu erwarten sind.

### 4.1. Optimierung der Dialoglevel-Bestimmung

Zur Beantwortung der Frage, wie der Schwankungsbereich weiter verkleinert werden kann, werden die am meisten abweichenden Filme noch einmal genauer untersucht. „Lissi und der wilde Kaiser“ würde nach einer wie oben beschriebenen Dialog-Normalisierung mit -3,4 dB zu leise wiedergegeben, er hat die größte negative Abweichung und wurde bereits aufgrund seiner kompakten Mischung besprochen (Abbildung 14). Interessant ist, dass der Film mit 65% auch den größten Sprachanteil in der Datenbank besitzt. „R U There“ weist mit 17% den geringsten Sprachanteil auf und hat mit +3,9 LU auch die größte positive Abweichung. Eine überzeugende Wechselwirkung der beiden Größen liegt jedoch nicht vor, wenn man andere Filme mit besonders hohen oder niedrigen Sprachanteilen dazu vergleicht. Dennoch ist die Information nicht wertlos: Wenn wie bei „Lissi und der wilde Kaiser“ in 65% der Zeit Dialog erkannt wird, dann ist klar, dass es sich hierbei um eine Komödie handelt und dass die Sprache das zentrale Element ist. Die Dialoge in solchen Filmen zeichnen sich nicht selten durch eine Sprachsynchon-Charakteristik aus, die im besten Fall ähnlich wie bei Werbung einen breiteren und bei gleicher Lautstärke auch teilweise angenehmeren Klang hat als Originalton-Dialoge. Die Authentizität vom Originalton wird bewusst durch die Prägnanz vom ADR ersetzt. Bereits im Kapitel 3.5.1. wurde festgestellt, dass Komödien dieser Art im Vergleich zu anderen Genres ein gesteigertes Dialoglevel aufweisen, selbst wenn sie von gleichen Mischtonmeistern gemischt wurden. Auch „La Clinique de l'Amour“ würde nach der Normalisierung mit 1,9 LU, „Traumschiff Surprise“ mit 2,7 LU zu leise wiedergegeben. Das heißt also, dass diese Filme nicht nur mit einem geringfügig lauterem Dialog gemischt werden, sondern dass sogar beim Publikum eine solche Erwartungshaltung entsteht. In der Sprache liegen die Jokes und Pointen;



das filmemacherische Kalkül bei solchen Komödien scheint zu sein, diese zu pushen, auch wenn oder gerade weil sich der Dialog dann ein Stück weit aus seinem natürlichen Kontext bewegt. Unabhängig davon, ob ein solches Phänomen als positive Strömung zu werten ist, kann es nach einer Dialog-Normalisierung selbstverständlich nicht erhalten bleiben.

Bei „R U There“, dem Film mit der größten positiven Abweichung, liegt ein ganz anderes Problem vor. Der niedrige Sprachanteil von 17% lässt bereits darauf schließen, dass für die Bestimmung des Dialoglevels nur wenige Dialogpassagen erkannt wurden, wodurch Fehler stärker ins Gewicht fallen. Der Verlauf des ermittelten Dialoglevels fällt im Verlauf des Films auf untypische Weise ab. Nach dem ersten Viertel liegt das Dialoglevel noch bei -30 LKFS, erfährt dann durch besonders leise Dialogstellen in der Mitte des Films zwei deutliche Einbrüche und sinkt bis zum Ende des Films auf -33,2 LKFS ab.

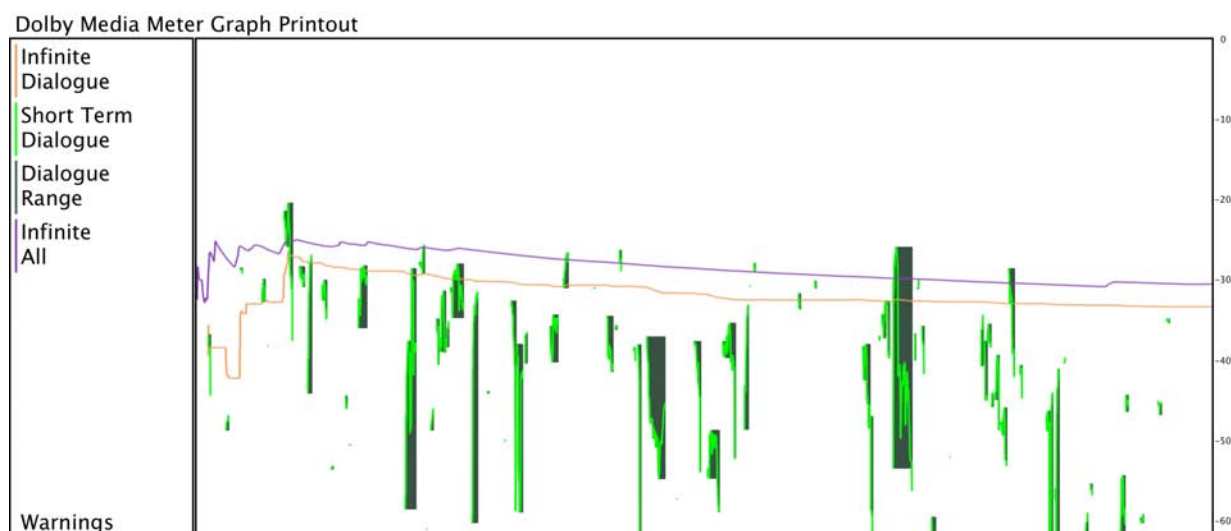


Abbildung 19: Entwicklung des Dialoglevels nach BS.1770-1 im Film „R U There“

Solche unsteten Mess-Ergebnisse sind problematisch, weil dann genau wie bei der Programmlautheit verschiedene Teile des Films verschiedene Werte liefern. Würde man die einzelnen Rollen dieses Films nach Dialog normalisieren, kämen bereits Sprünge zustande. Die Stärke der Dialoglevel-Normalisierung liegt aber gerade darin, dass ein Wert stellvertretend für den ganzen Film gilt.

Untersuchungen ergaben, dass einige der sehr leisen Dialogpassagen in „R U There“ und vor allem die zwei Stellen, welche die Dialoglevel-Einbrüche in der Mitte des Filmes hervorrufen, gar keinen Dialog enthalten. An diesen Stellen sind über weite Strecken z.T. nur Grillen und Vogellaute sowie Bewegungsrascheln und Computer Tippen vorhanden. Was im Einzelnen die fehlerhafte bzw. falsch-positive

Spracherkennung von Dialogue Intelligence hervorruft bleibt unklar. Fest steht jedoch, dass solche akustischen Elemente, deren Lautstärke weit unterhalb des durchschnittlichen Dialoglevels liegt, nicht mit in die Messung einfließen sollten. Ebenso gut könnten in Filmen leise Hintergrundgespräche, Lautsprecher-Durchsagen an Flughäfen, Radiosendungen im Nachbarzimmer als Dialogpassagen gewertet werden, wenngleich diese gar keinen szenischen Dialog darstellen. Solche Elemente führen dann dazu, dass das Dialoglevel geringer beziffert wird, als es eigentlich ist; der Film würde nach einer Normalisierung lauter wiedergegeben. In der Kinomischung sollte es keine Einschränkungen geben, wie leise bestimmte Sprachelemente eingesetzt werden. Während man natürliche Gespräche automatisch auf einem typischen Dialoglevel halten würde, muss es erlaubt sein, ausgewählte Stellen wesentlich leiser zu mischen, ohne dass dadurch gleich der ganze Film lauter wird. Ziel einer Lautstärke-Normalisierung kann es ja nur sein, kategorisch also dauerhaft zu laut gemischte Dialoge zu erkennen und dementsprechend zu korrigieren. Die beschriebene Methode soll also durch ein Gate erweitert werden, um stark vom Durchschnitt abweichende Dialoge von der Messung auszuschließen.

#### ***4.1.1. Implementierung eines relativen Gates***

Da auf dem Markt keine Implementierungen von Dialogue Intelligence in Verbindung mit relativen Lauthheits-Gates existieren, wurden in dieser Arbeit verschiedene Gates simuliert und empirisch ermittelt, welches sich am besten eignet. Da das Dolby Media Meter 2 keine Rohdaten zur Dialog-Lautheit ausgibt, wurden zur Simulation die Short-Term Dialogue Werte als Grundlage verwendet, welche über 3 Sekunden gefenstert sind. Dadurch entsteht ein Fehler, der nach Abschluss der Simulation kompensiert wurde.<sup>1</sup>

Gegen ein einfaches Gate bei -10 LU (jeweils relativ zum ohne Gate generierten Dialoglevel), ein Doppel-Gate bei +10 und -10 LU, ein Doppel-Gate bei +8 und -8 LU, ein Doppel-Gate bei +6 und -10 LU, sowie eine logarithmische Mittelung zwischen +10 und -10 LU hat sich ein einfaches Gate bei -8 LU durchgesetzt. Dadurch kann der Schwankungsbereich bis auf 5,3 LU gesenkt werden (Tabelle 6). Die Messungen für Filme, deren Dialoge besonders kompakt gemischt wurden, wie z.B. die Komödien „Chéri“, „Le Code a Changé“, „Lissi und der wilde Kaiser“ sowie „Blutzbrüdad“, werden durch das so konzipierte Gate fast gar nicht beeinflusst; ihre Dialoglevel fallen mit +0,1 LU marginal größer aus. Die Dialoge in „R U There“ hingegen werden mit +2,1 LU lauter und somit treffender gewertet. Falsch erkannte Dialogpassagen

weit unterhalb des szenischen Dialogs, werden aus der Wertung ausgeschlossen und ermöglichen so ein konsistenteres Ergebnis. Das Gate dient also sowohl zur Beibehaltung der freien künstlerischen Gestaltung als auch zur zusätzlichen effektiven Fehlerkorrektur des Dialogue Intelligence Algorithmus. Es verhindert zudem, dass das über den ganzen Film gemessene Dialoglevel durch zwischenzeitlich leise gemischte Dialogpassagen bewusst gedrückt und somit manipuliert wird, um nach einer Normalisierung lautere Wiedergaben zu bewirken.

Originaltitel	Dialoglevel 1770-1 mit rel. Gate bei -8 LU (LKFS)	Result. Dialoglevel 1770-1 mit rel. Gate bei -8 LU (LKFS) nach subj. Pegelung	Erzielte Änderung durch rel. Gate bei -8 LU (LU)	Abweichung vom Mittelwert nach Normalisierung
3	-29,6	-33,6	0,3	2,0
8 Femmes	-23,7	-30,7	0,2	-1,0
Alexander	-22,9	-28,9	0,8	-2,8
Alle Anderen	-29,4	-33,4	0,5	1,7
Another Year	-29,8	-33,8	0,6	2,1
Basic Instinct 2	-28,9	-31,9	0,2	0,2
Biutiful	-21,7	-29,7	1,2	-2,0
Blutzbrüdadz	-24,7	-29,7	0,3	-2,0
Caotica Ana	-29,2	-33,2	0,4	1,6
Chéri	-30,9	-33,9	0,1	2,2
Deutschland 09	-29,1	-33,1	0,3	1,4
Die Bluthochzeit	-28,4	-33,4	0,5	1,7
Elementarteilchen	-28,8	-32,8	0,6	1,2
Endlich	-30,2	-33,2	0,1	1,6
Erkan & Stefan 2	-24,3	-31,3	0,0	-0,4
Glück	-29,1	-33,1	0,9	1,4
Gomorrha	-25,7	-31,7	0,8	0,0
Hachi: A Dog's Tale	-31,0	-32,0	0,5	0,3
Hallam Foe	-27,1	-32,1	1,0	0,4
Heartbreak Hotel	-28,1	-33,1	0,4	1,4
Immortals	-17,8	-28,8	2,0	-2,9
Kammerflimmern	-35,7	-33,7	0,7	2,0
La Clinique de l'Amour!	-26,2	-30,2	0,2	-1,5
La Mome	-22,9	-28,9	0,6	-2,7
Le Code a Changé	-30,1	-33,1	0,1	1,4
Lian ai zhong de Bao Bei	-27,5	-32,5	1,6	0,8
Lissi und der wilde Kaiser	-20,8	-28,8	0,1	-2,8
Little Thirteen	-28,0	-32,0	1,1	0,3
Lost in Translation	-27,7	-31,7	1,5	0,0
Michael Clayton (DUB)	-28,8	-29,8	0,2	-1,9
Michael Clayton (OV)	-29,6	-30,6	0,2	-1,1
Mine Vaganti	-27,6	-31,6	0,3	-0,1
Ödland	-29,5	-32,5	0,4	0,8
Pandorum	-23,9	-31,9	0,6	0,2
Perfume - Story of a Murderer	-25,1	-31,1	0,8	-0,6
Precious	-29,6	-32,6	0,3	0,9
R U There	-31,1	-34,1	2,1	2,4
Rammbock	-27,0	-32,0	1,4	0,3
Resident Evil 3: Extinction	-19,9	-28,9	1,6	-2,8
Step Up 3D	-24,8	-29,8	0,6	-1,9
Step Up Revolution	-25,0	-30,0	0,4	-1,7
Tamara Drewe	-29,0	-33,0	0,3	1,3
The Dark	-25,2	-29,2	2,7	-2,5
The International	-28,8	-33,8	0,3	2,1
The Passion of the Christ	-25,0	-31,0	0,8	-0,7
The Three Musketeers (DUB)	-25,2	-33,2	0,1	1,5
The Three Musketeers (OV)	-25,1	-33,1	0,5	1,4
Traumschiff Surprise	-24,3	-29,3	0,3	-2,3
Türkisch für Anfänger	-27,3	-31,3	0,3	-0,4
Vorstadtkrokodile	-26,2	-31,2	0,4	-0,5
	<b>Mittelwert</b>	<b>-31,7</b>	Größte positive Abweichung	2,4
			Größte negative Abweichung	-2,9
			<b>Schwankungsbereich</b>	<b>5,3</b>

Tabelle 6: Auswirkung des relativen Gates bei -8 LU

## 4.2. Anpassung des Lautheits-Algorithmus

Der im DMM2 verwendete BS.1770-1 Algorithmus ist auf Rundfunk-Inhalte ausgerichtet und müsste für eine Kinoanwendung zwangsläufig angepasst werden. Weder die Kurve X, noch die Einmessung der 5.1-Kanäle im Kino werden derzeit berücksichtigt. Zudem erscheint es denkbar, dass tief-frequente Anteile, welche in einer TV-Abhörsituation kaum ins Gewicht fallen und zu Recht durch die K-Gewichtung aus der Messung nach BS.1770 ausgeschlossen werden, im Kino mehr zur Lautstärke-Wahrnehmung beitragen. Vor allem ist damit natürlich auch der LFE-Kanal gemeint, welcher bislang kategorisch aus der Messung ausgeschlossen wird und im Kino beachtliche Pegel erzeugen kann. Sicherlich sind solche tiefen Frequenzen, auch wenn sie sehr laut auftreten, nur selten als schmerzhaft laut einzustufen, es sollte dennoch untersucht werden, welchen Stellenwert sie bei der Lautstärke-Empfindung einnehmen und in welchem Maße der LFE in die Messung mit einfließen sollte. Zudem wird beim Lautheits-Algorithmus der ITU nicht berücksichtigt, dass die Surround-Kanäle im Kino seitens der Wiedergabe um 3 dB abgesenkt werden. Die K-Gewichtung müsste auch durch die im Kino übliche Kurve X erweitert werden, welche Tiefen unter 63 Hz sowie Höhen ab 2 kHz zunehmend absenkt. Da diese Absenkung von der Saalgröße des Kinos abhängt, muss bei der Filterung zwangsläufig ein Kompromiss getroffen werden.

## 4.3. Nomenklatur

Während „Dialoglevel“ durch die Bezeichnung „Level“ (dt. „Pegel“) den Anschein einer physikalischen Größe macht, sollte fortan die Bezeichnung „Dialog Loudness“ (dt. „Dialoglautheit“) benutzt werden, da es sich um eine psychoakustische Größe handelt.

Jegliche Anpassungen des Lautheits-Algorithmus an den Kino-Standard nach SMPTE RP200 sollten zudem als Abweichungen vom ITU-Standard durch die Bezeichnungen und Einheiten der Parameter kenntlich gemacht werden. Der Verweis „Theatrical“ mit dem Kürzel T bietet sich an, da M für „Movie“ bereits Dolbys  $L_{EQ(M)}$ -Verfahren kennzeichnet und C für „Cinema“ mit der C-Gewichtung für hohe Lautstärkepegel verwechselt werden könnte. Sinnvoll wäre daher die Bezeichnung „Theatrical Integrated Dialog Loudness“ sowie eine Referenzierung der absoluten Einheit als  $L_{UFS_T}$ . Denkbar wäre alternativ auch die Einheit  $L_{TFS}$  (Loudness, T-Weighting, referenced to digital Full Scale), je nachdem ob eine Einigung auf  $L_{UFS}$  (ISO 80000-8) stattfindet oder nicht. Die relative „Loudness Unit“

in LU sowie das „True Peak Level“ in dBTP zur technischen Aussteuerung bleiben selbstverständlich unverändert. Vervollständigt werden diese Parameter dann durch „Theatrical Program Loudness“, „Theatrical Short-Term Loudness“, „Theatrical Momentary Loudness“, „Theatrical Short-Term Dialog Loudness“ (zur Aussteuerung des Dialogs in Echtzeit) sowie „Theatrical Loudness Range“ und ggf. „Theatrical Dialog Range“ (Kapitel 3.5.3.).

#### **4.4. Maximale Lautheit**

Zusätzlich zur Dialog-Normalisierung muss untersucht werden welche Spitzen-Lautstärken nicht überschritten werden sollten und wie diese durch die Parameter Short-Term und Momentary Loudness beschrieben werden. Hierzu wurden die bei den subjektiven Hörtests des Autors erstellten Anmerkungen zu besonders lauten Stellen einiger Filme ausgewertet (Tabelle 7). Die Filmausschnitte sind unterteilt in anhaltend laute Passagen, die zur Beurteilung der Short-Term Loudness Werte zur Verfügung stehen, sowie kürzere Lautheits-Spitzen zur Beurteilung der Momentary Loudness Werte. Neben der Beschreibung zum Inhalt ist auch vermerkt, wenn die Elemente einen besonders hohen LFE- bzw. Surround-Anteil haben. Dass solche beim ITU-Algorithmus nicht gewertet bzw. falsch gewichtet werden, wird bei der Analyse der ermittelten Werten durchaus erkennbar:

Bei der Short-Term Loudness zeichnet sich nach Begutachtung der Daten die Tendenz ab, dass bei -12 LUFS eine Toleranzgrenze erreicht wird. Die Passagen, die noch am meisten davon abweichen, sind Musiksequenzen mit viel LFE-Anteil. Würde der Lautheits-Algorithmus diese stärker berücksichtigen, wären auch die zugehörigen Short-Term Loudness Werte höher und somit näher bei -12 LUFS. Bei der Momentary Loudness hingegen scheinen maximal -8 LUFS akzeptabel zu sein, und auch hier gibt es Abweichungen nach unten bei besonders tieffrequenten Elementen. Außerdem wird die Grenze von -8 LUFS zweifach überschritten und zwar von zwei kurzen Effekten, die jeweils einen großen Surround-Anteil haben, wodurch fälschlicherweise größere Lautheits-Werte entstehen: Zum einen handelt es sich um einen in den Surrounds brüllenden Elefanten in „Alexander“ (die lauteste Stelle im Film), zum anderen um einen Schock-Effekt in „Immortals“, der auf allen 5.0-Kanälen gleichmäßig laut gemischt ist. Wegen der in BS.1770 definierten erhöhten Gewichtung der Surround-Kanäle, welche sich aber im Kino durch die 3 dB Absenkung so gar nicht überträgt, würde eine angepasste Messung niedriger ausfallen.

Originaltitel	TC	Beschreibung	subj. Pegelung Autor	Max. S.-Term Loudness (LUFS)	Max. Mom. Loudness (LUFS)	Max. S.-Term L. nach subj. Pegelung (LUFS)	Max Mom. L. nach subj. Pegelung (LUFS)	Max. S.-Term L. bei auf -8 norm. Max Mom. L. (LUFS)	Max. Mom. L.. Bei auf -12 norm. Max. S.-Term L. (LUFS)
Alexander	01:48:36	Musik und Crowd-Jubeln	-6	-5,4	-4,0	<b>-11,4</b>	-10,0	-9,4	-10,6
Blutzbrüdad	02:18:51	Musiksequenz m. Rap	-4	-9,8	-8,6	<b>-13,8</b>	-12,6	-9,2	-10,8
Gomorrha	12:08:10	Abspannmusik steigend	-6	-5,7	-4,1	<b>-11,7</b>	-10,1	-9,6	-10,4
Perfume - Story of a Murderer	01:02:55	Crowds Rufe	-6	-7,0	-4,1	<b>-13,0</b>	-10,1	-10,9	-9,1
Precious	01:06:48	FX breit, Übergang zu Musik	-2	-15,1	-12,3	<b>-17,1</b>	-14,3	-10,8	-9,2
Resident Evil 3: Extinction	02:24:23	Schreien und FX	-8	-3,9	-1,6	<b>-11,9</b>	-9,6	-10,3	-9,7
Step Up 3D	01:48:17	Musiksequenz	-5	-9,2	-4,8	<b>-14,2</b>	-9,8	-12,4	-7,6
Step Up Revolution	01:10:27	Flame FX Musiksequenz m. lauten Crowds	-6	-8,5	-7,2	<b>-14,5</b>	-13,2	-9,3	-10,7
Step Up Revolution	02:24:50	Musiksequenz FX und Crowd-Schreien, viel Bass	-7	-8,2	-5,9	<b>-15,2</b>	-12,9	-10,3	-9,7
Tamara Drewe	10:29:54	Rückkopplung auf Rednerpult, mit Surround	-5	-6,8	-2,7	<b>-11,8</b>	-7,7	-12,1	-7,9
Tamara Drewe	11:33:06	Vorbeiziehende Herden	-5	-7,4	-3,0	<b>-12,4</b>	-8,0	-12,4	-7,6
The Dark	02:14:59	Sequenz mit Schreien und FX	-4	-8,1	-6,1	<b>-12,1</b>	-10,1	-10,0	-10,0
The Three Musketeers (OV)	02:12:03	Schlachtsequenz	-6	-7,4	-5,1	<b>-13,4</b>	-11,1	-10,3	-9,7
The International	11:18:11	Kronleuchter	-7	-4,6	-2,4	<b>-11,6</b>	-9,4	-10,2	-9,8
		<b>Kurze Lautheits-Spitzen:</b>							
Alexander	03:14:59	Elefantenschrei, vor allem Surround	-8	-2,1	1,5	-10,1	<b>-6,5</b>	-11,6	-8,4
Gomorrha	11:32:37	Pistolenschuss	-6	-11,0	-6,0	-17,0	<b>-12,0</b>	-13,0	-7,0
Immortals	02:27:35	2x "Amboss" FX, viel Bass	-11	-2,7	1,0	-13,7	<b>-10,0</b>	-11,7	-8,3
Immortals	02:28:00	FX, breitbandig, viel Surround	-11	-1,3	4,5	-12,3	<b>-6,5</b>	-13,8	-6,2
Pandorum	01:36:57	hochfrequentes Zwischen	-9	-1,6	0,2	-10,6	<b>-8,8</b>	-9,8	-10,2
Perfume - Story of a Murderer	01:05:42	sehr bassiger FX	-5	-13,8	-8,8	-18,8	<b>-13,8</b>	-13,0	-7,0
Precious	02:17:04	Musiksequenz mit transientenreichen FX	-2	-14,0	-10,8	-16,0	<b>-12,8</b>	-11,2	-8,8
Resident Evil 3: Extinction	02:21:31	FX schrill	-9	-4,6	-1,4	-13,6	<b>-10,4</b>	-11,2	-8,8
Step Up Revolution	02:24:34	Basslastiger FX, Musiksequenz, Crowd-Schreien	-6	-7,7	-5,9	-13,7	<b>-11,9</b>	-9,8	-10,2
The Dark	01:27:05	Effekt breitbandig, viel Bass, viel Surround	-3	-10,9	-5,4	-13,9	<b>-8,4</b>	-13,5	-6,5
The Dark	02:01:52	Schrei	-6	-8,1	-3,4	-14,1	<b>-9,4</b>	-12,7	-7,3
The International	10:39:11	Schüsse	-6	-11,2	-7,0	-17,2	<b>-13,0</b>	-12,2	-7,8
The International	11:17:01	Kronleuchter kurz	-9	-5,9	-3,5	-14,9	<b>-12,5</b>	-10,4	-9,6

Tabelle 7: Short-Term und Momentary Loudness besonders lauter Filmszenen

Abschließend muss geklärt werden, ob wirklich zwei Lautheits-Obergrenzen notwendig sind, oder ob eine im Bezug zur anderen redundant ist. Werden zunächst alle Passagen auf eine Momentary Loudness von -8 LUFS normalisiert, dann wird die Short-Term Loudness von -12 LUFS noch häufig deutlich überschritten. Wird hingegen auf letztere normalisiert, dann resultieren Momentary Loudness Werte, die -8 LUFS nur marginal überschreiten. Die zwei Filme, welche mit mehr als 1 LU abweichen würden, besitzen ihrerseits wieder einen hohen Surround-Anteil, der durch einen angepassten Lautheits-Algorithmus weniger berücksichtigt würde. Dadurch ist anzunehmen, dass für die Modellnorm zur Kinomischung eine Obergrenze für Short-Term Loudness ausreicht, da die Grenze zur Momentary Loudness automatisch eingehalten wird. Eine solche Regelung bewirkt auch, dass besonders kurze laute Geräusche wie Schüsse einzeln eingesetzt sehr laut gemischt werden können. Das gleiche Geräusch mehrfach wiederholt würde aber die Short-Term Loudness Grenze überschreiten. Gäbe es hingegen nur eine Momentary Loudness Grenze, dann würde es weniger eine Rolle spielen, ob das laute Geräusch einzeln vorkommt oder mehrfach in kurzen Abständen wiederholt, da es in der Messung immer wieder von einem neuen Fenster erfasst würde. Die Charakteristik der Short-Term Loudness hat also in diesem Kontext auch einen höheren Mehrwert, da sie einen dynamischeren Umgang bei besonders lauten Stellen im Film anregt.

#### **4.5. Festlegung der Zielwerte**

Wenn nun die erarbeitete Dialog-Normalisierung zur Modellnorm werden soll, dann muss ein konkreter Zielwert für Dialoge definiert werden. Den vorliegenden Ergebnissen zufolge liegt dieser Wert, also die angestrebte Theatrical Integrated Dialog Loudness, bei -31,7 LKFS, folglich der Ideallautstärke des Autors. Wie im 1. Kapitel besprochen ist die Ideallautstärke von Person zu Person unterschiedlich, sie kann im Kino ohnehin nicht frei gewählt werden. Die Lautstärke sollte von Film zu Film jedoch kontinuierlich sein, sodass Kinos bzw. Filmvorführer sich dann mit definierten Wiedergabe-Levels am Kinoprozessor an Zielgruppen und deren Lautstärke-Erwartungen richten können. Der Zielwert könnte also frei gewählt werden, er entscheidet lediglich, welche Lautstärke Filmvorführer im Bezug zur Dolby Skala zu erwarten hätten und wo der Arbeitspunkt auf der digitalen Aussteuerung liegt. Wichtig ist aber, dass dabei eine möglichst hohe Korrelation zum SMPTE Standard RP200 entsteht, dass also bei Referenz-Level 7 die „richtige“, von der Produktion angestrebte Wiedergabe erreicht wird. In der Praxis hat sich gezeigt,

dass Filme, selbst wenn sie auf Referenz-Lautstärke gemischt wurden, bei Level 7 laut sein sollten, d.h. tendenziell lauter als nötig, ohne dabei jedoch wehzutun. Für Vorführer besteht bei dieser Marke ohnehin eine Art Schallmauer, die womöglich auch dann lange nicht überschritten würde, wenn wieder leiser gemischt würde. Es macht also Sinn die dieser Arbeit zu Grunde liegende Ideallautstärke knapp unter Level 7 anzusiedeln und den Zielwert für Dialoge nicht bei -31,7 LKFS sondern -30,0 LKFS festzulegen. Damit liegt die in dieser Arbeit angenommene Ideallautstärke exakt auf Dolby Level 6,5. Level 7 steht dadurch weiterhin für eine kräftige Wiedergabe.

Das Ziel dieser Modellnorm besteht darin, übermäßig laute Mischungen, die auf zu leisen Abhörlautstärken erstellt wurden, zu erkennen und zu verhindern. Es liegt in der Natur der Sache, dass diese Mischungen eine erhöhte Dialog Loudness aufweisen und die gesamte Kino-Verwertungskette wie im 1. Kapitel beschrieben negativ beeinflussen. Das bedeutet jedoch nicht, dass Mischungen, die zu leise erstellt wurden, sich ähnlich negativ auf das System auswirken; sie sollten deshalb nicht aufwärts normalisiert werden! Es liegt in der Hand der Filmemacher Dialoge nach Belieben leiser zu mischen und stattdessen andere Elemente in den Vordergrund zu stellen. Solange die Filme auf Referenzlevel gemischt werden und richtig klingen, ist auch nicht zu erwarten, dass Vorführer in solchen Ausnahmefällen gleich die Kinoprozessoren lauter stellen, nur weil die Dialoge leiser sind. Ausnahmen sind nämlich auch Filme, die zu wenig oder gar keinen Dialog haben. Zwar sind solche Filme heutzutage sehr selten, jedoch besteht natürlich der Verdacht, dass diese dann stark abweichende integrierte Dialoglautheiten aufweisen, da zu wenig oder gar keine Dialogpassagen für die Messung zur Verfügung stehen. Aus diesem Grund wurden nochmals zwei Pseudo-Stummfilme „The Artist“ (2011, Regie: Michel Hazanavicius) und „Silent Movie“ (1976, Regie: Mel Brooks) sowie ein Film mit besonders wenig Dialog „The Turin Horse“ (2011, Regie: Béla Tarr) anhand ihrer kommerziell erwerblichen DVDs gemessen. „The Artist“ scheint durch seinen Dynamikverlauf als Einziger in seiner Original Kinofassung vorzuliegen, während die Tonspuren zu „Silent Movie“ und „The Turin Horse“ für die DVDs anscheinend komprimiert wurden. Die Filme kommen auf sehr kleine Sprachanteile von 3% bis 8% und weisen mit integrierten Dialoglautheiten von -30,9 LKFS („The Artist“), -27,4 LKFS („Silent Movie“) und -32,2 LKFS („The Turin Horse“) erstaunlich gewöhnliche Werte auf. Wenngleich es hier zahlreiche Dialogue Intelligence Fehldetektionen gibt, werden üblicherweise die wenigen doch vorhandenen Dialoge gewertet (z.B. eine



Gruppe lachender Schauspielerinnen in einer kurzen vertonten Sequenz in „The Artist“). Außerdem fallen bei Stummfilmen besonders häufig Musikelemente in die Dialogwertung, wenn diese Sprache imitieren oder karikieren (z.B. Posaune mit Dämpfer). Demnach würden „The Artist“ und „The Turin Horse“ durch die Modell Norm nicht aufwärts normalisiert und in ihrer ursprünglich gemischten Form wiedergegeben. „Silent Movie“ müsste um 2,6 LU abgesenkt werden, was nach subjektiver Einschätzung im Kino als richtig zu werten ist. Es mag jedoch theoretisch noch extremere Beispiele geben, die sehr geringe integrierte Dialoglautheiten hervorrufen könnten. Man nehme z.B. an, ein Film enthält keinen szenischen Dialog und keine Musik, sondern nur Hintergrund-Atmos oder noch extremer nur technisches Rauschen. Es wäre fatal, wenn solche Filme dann um 30 LU oder mehr verstärkt würden. Sie sollten zwangsläufig in ihrer Ursprungsform belassen werden, weil sie durch ihr „Zu-leise-sein“ die Projektionen anderer Filme im selben Kino nicht negativ beeinflussen wie es zu laute Filme tun. Filme ohne Dialog hätten nach diesem Modell aber auch die Möglichkeit, wesentlich lauter gemischt zu werden, wenn gemessene Dialoglautheiten sehr niedrig ausfallen. Sollte sich diese Tatsache - und es geht dabei wohlgerne um einen verschwindend geringen Prozentsatz gegenwärtiger Kinoproduktionen - nachhaltig negativ auswirken, sollte evaluiert werden, die vorgestellte Modellnorm durch eine Obergrenze zur Programmlautheit zu ergänzen, welche die integrierte Dialoglautheit ersetzt, sobald der Sprachanteil eine definierte Marke (z.B. 10%) unterschreitet. Die Funktionsweise einer solchen Schaltung wurde bereits in Kapitel 3.5.3. erläutert (Abbildung 17).

Die Theatrical Short-Term Loudness Obergrenze wird ebenfalls wie oben beschrieben um 1,7 LU angehoben und aufgerundet; sie liegt dann bei -10 LUFs. Dieser Wert wird nicht nur durch die subjektiven Hörtests in dieser Arbeit sondern auch durch die Tatsache bekräftigt, dass dadurch über dem Dialog wieder ein natürlicher Loudness Headroom von 20 LU entsteht. Dieser entspricht wiederum dem technischen Headroom der SMPTE Norm RP200, welche das Referenzlevel bei 20 dB unter digitaler Volllaussteuerung  $0\text{dB}_{\text{FS}}$  definiert.

## 4.6. Umsetzung

Denkbar sind zwei Anwendungsformen der Modellnorm: Zum einen könnte sie Abhör-seitig in digitalen Kinos angewendet werden. Dazu muss vor Ort die Dialog-Normalisierung auf den Zielwert sowie eine Absenkung bei Überschreitung der Short-Term Loudness Obergrenze stattfinden. Die Eingriffe erfordern als Grundlage

Lautheitsmessungen sowohl zur maximalen Theatrical Short-Term Loudness sowie zur Theatrical Integrated Dialog Loudness gemessen mit Dialogue Intelligence und relativem Gate bei -8 LU. Die Messungen sollen nicht in Echtzeit während der Wiedergabe stattfinden, sondern vor der Vorführung durchgeführt und abgeschlossen werden, sodass die Filme von Beginn an korrekt gezeigt werden. Für das Gate von -8 LU bedeutet das, dass anders als beim relativen Gate nach BS.1770-2 keine Echtzeit Implementierung notwendig ist; die Gate-Korrektur kann nach abgeschlossener Dialoglautheit-Bestimmung stattfinden (dann also wenn bereits ein stabiler Dialoglautheits-Wert feststeht, in dessen Relation das Gate arbeiten kann).

Ein Prozessor, der solche Lautheits-Anpassungen bewerkstelligen kann, wurde mit dem Program Optimizer DP600 für den Rundfunk-Bereich bereits von Dolby selbst auf den Markt gebracht. Mit diesem können Dialogue Intelligence Messungen nach BS.1770-1 und BS.1770-2 durchgeführt werden, wobei selbst MXF, das Audio-Hüllformat im DCP unterstützt wird. Eine solche Einheit müsste wie in Kapitel 4.2. beschrieben für den Kino-Einsatz angepasst, durch ein relatives Gate erweitert und auf den entsprechenden Zielwert von -30 LKFS eingestellt werden. Anschließend würde sie zwischen Digital Cinema Server und dem Kinoprozessor eingeschliffen, sodass die Korrektur bereits vor dem Dolby Regler stattfindet. Es reicht allerdings nicht aus, wenn der Prozessor den Audio-Stream in Echtzeit während der Vorführung vom Server erhält, sondern er muss bereits im Voraus File-basiert darauf zugreifen können, um die Messungen durchzuführen. Ein solcher Zugriff auf den Server ist bislang nicht vorgesehen und er steht insbesondere im Konflikt zur KDM (Key Delivery Message) bei encrypteten Kopien, welche die DCP Daten erst dann entschlüsselt, wenn die Vorführung startet. Die hier erläuterten Vorschläge greifen also weit in die bestehende Funktionsweise des digitalen Kinos ein. Es ist außerdem zu erwarten, dass eine solche Technologie nur teilweise von Kinos implementiert würde; sie kann dem Markt nicht aufgezwungen werden. Die beschriebene Methode würde für alle DCPs funktionieren, also auch für solche, die vor dem Zeitpunkt der Umstellung erstellt wurden, nicht jedoch für 35mm Kopien, da hier keine schnellen Lautheits-Messungen durchgeführt werden könnten.

Eine zweite Anwendungsform wäre, dass die beschriebene Modellnorm zur Mischungs-Vorschrift wird. Idealerweise müsste die DCI (Digital Cinema Initiatives) dafür den DCP-Standard durch entsprechende Vorgaben zu den Audio Characteristics erweitern. Da die Organisation bereits ein Zusammenschluss der größten amerikanischen Film-Studios ist und sich ähnliche Vorschriften wie etwa die

der TASA flächendeckend in Europa durchsetzen konnten, wäre eine Umsetzung des Standards möglich.

Wenn in Zukunft gar keine 35mm-Kopien mehr erstellt und somit auch keine Dolby Print Master, d.h. keine Kontrollen und Abnahmen der Mischungen mehr durchgeführt werden, dann muss eine solche Vorschrift anderweitig überwacht werden. Denkbar wäre ein Zertifikate-Modell nach Vorbild der TASA oder wie in Deutschland eine Überwachung durch die Freiwillige Selbstkontrolle der Filmwirtschaft (FSK). Alle mit dem Mischungs-Standard konformen DCPs sollten durch ein entsprechendes Siegel gekennzeichnet sein, damit Filmvorführer absehen könnten, ob es sich um eine Tonspur nach alter oder neuer Norm handelt.

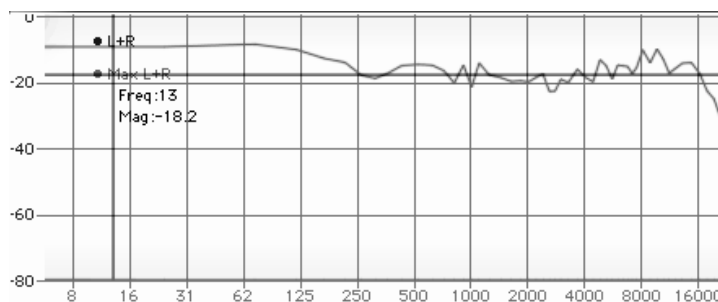
Die Umsetzung der Lautheits-Vorgaben würde dann genau wie bei TASA Trailern oder R 128 Formaten nicht bei der Wiedergabe sondern während der Mischung passieren. Zunächst muss dazu der Kinoprozessor im Studio wieder auf Referenzlevel gestellt werden. Allein dadurch wird sich die Dialoglautheit auf natürliche Weise um -30 LKFS bewegen, was je nach Charakteristik der Sprache, also z.B. Transienten-Reichtum und Komprimierungs-Grad etwa  $-20\text{dB}_{\text{FS}}$  (QPPM) entspricht, also dem seit jeher etablierten Dialoglevel. Gemischt werden kann und sollte weiterhin nach Gehör, wobei mit einem speziell für den Kino-Einsatz angepassten Lautheits-Messgerät oder alternativ Instanzen von LM100 oder DMM2 die Dialog Loudness und Short-Term Loudness sowie zur technischen Aussteuerung das True Peak Level im Blick gehalten werden sollten. Da im Gegensatz zu gegenwärtig üblichen Mischungen ein niedrigerer Arbeitspunkt gewählt wird, steht wieder der gesamte technische Headroom des digitalen Kinos zur Verfügung, was dynamische Mischungen ermöglicht. Während bei übermäßig laut gemischten Filmen mit erhöhtem Dialoglevel regelmäßig Limiter angefahren werden, um mehr Lautheit zu erreichen, kann hier auf exzessive Kompression verzichtet werden. Auch die Einhaltung der Short-Term Loudness Obergrenze ließe sich durch Limiter nicht sicher einhalten, da sich diese auf Spannungspegel und nicht Lautheits-Werte beziehen. Es reicht also nicht aus, laute Effekte lediglich gegen einen Limiter zu fahren, sie müssen gezielt angesteuert werden. Lediglich zum Schutz vor Clipping von Stems und Mastern sollten natürlich nach Belieben technische Limiter gesetzt werden. Nach Abschluss muss nochmals die über den gesamten Film integrierte und gegatete Dialoglautheit bestimmt werden. Sollte sie über  $-30\text{LUFS}_T$  liegen, muss die gesamte Mischung entsprechend abgesenkt werden, wodurch die Short-Term Loudness Obergrenze nicht überschritten werden kann, sofern sie zuvor eingehalten

wurde. Liegt sie hingegen unter  $-30 \text{ LUFS}_T$  kann die Mischung so belassen werden, da sie nach Gehör erstellt wurde und bei Referenzlevel richtig klingt.

## 4.7. Auswirkung

Um abzusehen wie stark die konzipierte Modellnorm in das bisherige System eingreifen würde und wie stark eine Lautheits-Obergrenze die bisherige Mischungsfreiheit einschränken würde, wurden ein weiteres Mal die Filme der Datenbank daraufhin untersucht, wie stark sie modifiziert werden müssten, um die festgelegten Anforderungen zu erfüllen. 44 der 50 Filme müssten in ihrer Gesamtheit mit zwischen 0,2 und 12,2 LU abgesenkt werden, um ein Dialoglevel von  $-30 \text{ LKFS}$  (gemessen nach ITU-R BS.1770-1, ohne Anpassung für Kino) zu erreichen. Sechs Filme bräuchten nicht korrigiert werden, da diese bereits ein Dialoglevel von  $-30 \text{ LKFS}$  oder weniger aufweisen. Während die Filme zuvor Short-Term Loudness Werte von bis zu  $+1,2 \text{ LUFS}$  erreichten, wird die Obergrenze von  $-10 \text{ LUFS}$  nach der Dialog-Normalisierung nur noch von vier Filmen überschritten. Der Elefant in „Alexander“ (2:14:59) müsste demnach um 1,2 LU abgesenkt werden, eine Effektsequenz in „Pandorum“ (0:36:57) um 2,3 LU, ein Mikrofon-Rückkopplungs-Effekt in „Tamara Drewe“ (0:29:54) um 2,2 LU. Alle diese Passagen weisen einen besonders starken Einsatz der Surround-Kanäle auf. Hier wäre bei einer für Kino angepassten Lautheits-Messung also ein etwas niedrigerer Short-Term Loudness Wert zu erwarten (Kapitel 4.2.). Im vierten Film „The International“ würde nach einer Dialog-Normalisierung die Short-Term Loudness Obergrenze mit ganzen 4,2 LU überschritten, wenn der Kronleuchter auf dem Boden des Guggenheim Museums zerschellt (1:18:11). Eine Frequenzanalyse zu dieser lautesten Stelle im Film (Abbildung 20) verdeutlicht jedoch, dass durch das Glas-Klirren besonders viel Energie in den hohen Frequenzen noch weit über 2 kHz besteht, welche im Kino durch die Kurve X weniger laut wahrgenommen würden. Zudem besteht auch bei dieser Passage ein nicht unwesentlicher Surround-Anteil. Untersuchungen mit entsprechender Filterung vor dem Messgerät ergaben, dass eine Erweiterung der K-Gewichtung durch eine Annäherung der Kurve X bei dieser Filmszene eine um etwa 1,5 LU geringere Short-Term Loudness bewirken würde. Eine um 3dB geringere Wertung der Surround-Kanäle würde bei den genannten Filmpassagen je nach Surround-Anteil zwischen 1,5 und 2,3 LU ausmachen. Da solche Änderungen jedoch Auswirkungen auf alle erfassten Daten hätten und somit auch möglicherweise auf die ermittelte Lautheits-Obergrenze, soll hier lediglich darauf

hingewiesen werden, dass die aufgeführten Überschreitungen voraussichtlich weniger drastisch ausfallen, wenn nicht ganz wegfallen würden. Fest steht vor allem, dass kaum einer der 50 Filme überhaupt durch die Short-Term Loudness Obergrenze eingeschränkt würde, die große Mehrheit bräuchte zusätzlich zur Dialog-Normalisierung nicht mehr korrigiert werden. Es sei auch angemerkt, dass solche Filme, die nach einer Normalisierung noch Spitzenwerte von -10 LUFS oder mehr erreichen, in diesem Kontext nicht exemplarisch für falsch und überkompensierend zu laut gemischte Produktionen stehen. Im Gegenteil wurden hier von vorneherein Dialoglevel gewählt, welche innerhalb des technischen Aussteuerungsbereiches



**Abbildung 20: Frequenzanalyse des fallenden Kronleuchters in „The International“**

überhaupt einen Lautheits-Headroom von 20 LU und somit dynamische Mischungen ermöglichen. Filme wie „Immortals“ erreichen nach einer Normalisierung nämlich auch deshalb keine -10 LUFS mehr, weil sie durch vermutlich zu niedrige Abhörlautstärken bei der Mischung viel näher an der digitalen Vollaussteuerung und somit komprimierter gemischt wurden.

Der Tatsache geschuldet, dass Filme mit Dialoglautheiten von unter -30 LKFS nicht aufwärts normalisiert werden, sondern auf ihrem ursprünglichen Wert belassen werden, entstehen letztendlich Abweichungen zu den in dieser Arbeit aufgestellten subjektiven Wiedergabe-Erwartungen mit einem endgültigen Schwankungsbereich von 5,8 LU. Wenngleich eine kategorische Normalisierung (d.h. auch aufwärts) mit 5,3 LU Schwankung marginal günstiger ausfallen würde, ist die hier definierte Methode, bei der nur zu laute Mischungen abgesenkt werden, unerlässlich (Kapitel 4.5.). Somit beträgt die endgültige größte positive Abweichung nur noch +2,1 LU, was sogar unter dem Wert von +2,4 dB liegt, der von TV-Zuschauern bezogen auf ihre Ideallautstärke als „louder but acceptable“ bezeichnet wurde. Die größte negative Abweichung liegt bei nunmehr -3,7 LU und damit noch deutlich über dem Wert „softer but acceptable“ von -5,4 dB (Abbildung 1, Kapitel 1.1.1.). Auf diese Weise liegen jetzt 100% der gemessenen Filme innerhalb der Komfortzone eines Zuschauers, vorausgesetzt er schaut diese auf einem für ihn passenden Dolby Level.

Originaltitel	Absolute Dialoglevel- Abweichung von der Modell Norm (LU)	Result. Dialoglevel nach Dialog Normalisierung (LKFS)	Result. Max. S-T. Loudness nach Dialog Normalisierung (LUFS)	Überschreitung der S-T. Loudness Obergrenze (LU)
3	0,4	-30,0	-14,6	
8 Femmes	6,3	-30,0	-18,0	
Alexander	7,1	-30,0	-8,8	1,2
Alle Anderen	0,6	-30,0	-16,9	
Another Year	0,2	-30,0	-19,1	
Basic Instinct 2	1,1	-30,0	-12,3	
Biutiful	8,4	-30,0	-16,7	
Blutzbrüdad	5,4	-30,0	-15,2	
Caotica Ana	0,8	-30,0	-11,1	
<b>Chéri</b>	<b>0,0</b>	<b>-30,9</b>	-14,8	
Deutschland 09	0,9	-30,0	-13,6	
Die Bluthochzeit	1,6	-30,0	-11,5	
Elementarteilchen	1,2	-30,0	-16,0	
<b>Endlich</b>	<b>0,0</b>	<b>-30,2</b>	-16,0	
Erkan & Stefan 2	5,7	-30,0	-11,0	
Glück	0,9	-30,0	-13,1	
Gomorrha	4,3	-30,0	-10,0	
<b>Hachi: A Dog's Tale</b>	<b>0,0</b>	<b>-31,0</b>	-15,6	
Hallam Foe	2,9	-30,0	-10,8	
Heartbreak Hotel	1,9	-30,0	-12,4	
Immortals	12,2	-30,0	-11,0	
<b>Kammerflimmern</b>	<b>0,0</b>	<b>-35,7</b>	-17,4	
La Clinique de l'Amour!	3,8	-30,0	-15,0	
La Mome	7,1	-30,0	-18,1	
<b>Le Code a Changé</b>	<b>0,0</b>	<b>-30,1</b>	-14,0	
Lian ai zhong de Bao Bei	2,5	-30,0	-14,5	
Lissi und der wilde Kaiser	9,2	-30,0	-14,0	
Little Thirteen	2,0	-30,0	-13,3	
Lost in Translation	2,3	-30,0	-10,1	
Michael Clayton (DUB)	1,2	-30,0	-16,0	
Michael Clayton (OV)	0,4	-30,0	-15,2	
Mine Vaganti	2,4	-30,0	-14,5	
Ödland	0,5	-30,0	-20,3	
Pandorum	6,1	-30,0	-7,7	2,3
Perfume - Story of a Murderer	4,9	-30,0	-11,9	
Precious	0,4	-30,0	-14,4	
<b>R U There</b>	<b>0,0</b>	<b>-31,1</b>	-12,6	
Rammbock	3,1	-30,0	-10,7	
Resident Evil 3: Extinction	10,1	-30,0	-14,0	
Step Up 3D	5,2	-30,0	-14,0	
Step Up Revolution	5,0	-30,0	-12,7	
Tamara Drewe	1,0	-30,0	-7,8	2,2
The Dark	4,8	-30,0	-12,9	
The International	1,2	-30,0	-5,8	4,2
The Passion of the Christ	5,1	-30,0	-13,6	
The Three Musketeers (DUB)	4,8	-30,0	-10,7	
The Three Musketeers (OV)	4,9	-30,0	-10,8	
Traumschiff Surprise	5,7	-30,0	-15,7	
Türkisch für Anfänger	2,7	-30,0	-11,5	
Vorstadtkrokodile	3,8	-30,0	-14,0	

Tabelle 8: Abweichungen von der Modellnorm (fett markierte Filme bleiben unverändert)

<sup>1</sup> Fehlerkorrektur bei Gate Simulation: Das Dialoglevel des jeweiligen Films wurde zunächst anhand der Short Term Dialogue - Werte berechnet, indem der Durchschnitt aller Lautheits-Werte von Dialogpassagen oberhalb von -70 LKFS (absolutes Gate) ermittelt und anschließend wieder logarithmisch dargestellt wurde. Die resultierende Abweichung zum vom DMM2 errechneten Dialoglevel (je nach Film zwischen +0,4 und -0,9 dB) wurde anschließend bei den Ergebnissen zum gegateten Dialoglevel kompensiert.

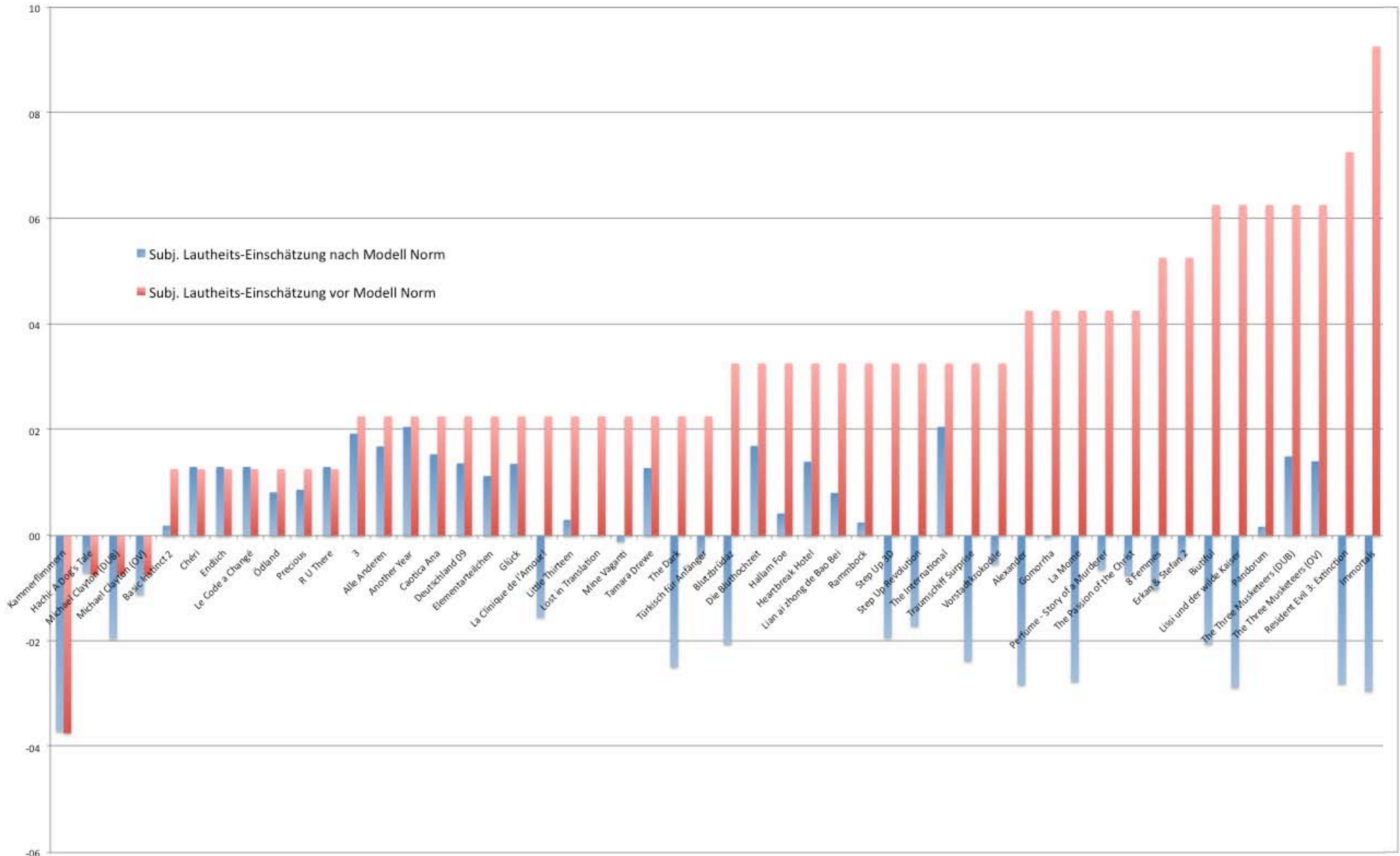


Abbildung 21: Finale Abweichungen von der Ideallautstärke des Autors vor und nach Normalisierung durch Modellnorm (Wiedergabe auf Dolby Level 6,5)

## Fazit

In dieser Arbeit wurde die aktuelle Lautstärke-Situation im vom SMPTE Standard RP200 geprägten Kino untersucht. Durch Analyse der Verhalten und Interessen aller beteiligten Parteien wurde befunden, dass das gegenwärtige System unzureichend ist und gefolgert, dass alle Filme sowie Trailer und Werbungen nach einer Norm gemischt werden müssen, durch die sie vom Publikum, unabhängig von Genre und dramaturgischer Dynamik, als gleich laut empfunden werden. Es wurde die These aufgestellt, dass selbst dann mit einem einzigen Wiedergabelevel nicht der gesamte Markt bedient werden kann, dass es daher möglich sein muss, für Kinder und jegliche andere Zielgruppen unterhalb der Referenzlautstärke zu projizieren und dass sich eine solche marktgerichtete Spartenbildung nur positiv für alle Konsumenten auswirken kann.

Drei wichtige etablierte Lautheits-Standards, die TASA Norm mit  $L_{EQ(M)}$ , die Dialog-Normalisierung von AC3 sowie die Lautheits-Normalisierung nach EBU R 128 und ATSC A/85 durch ITU-R BS.1770 wurden untersucht und anhand ihrer elf Parameter durch Messung von 50 Kinofilmen auf ihre Aussagekraft zur Lautstärke-Wahrnehmung im Kino geprüft. Mit den daraus gewonnenen Erkenntnissen wurde die folgende Modellnorm aufgestellt:

Ein Kinoprogramm soll primär durch seine „Theatrical Integrated Dialog Loudness“ in  $L_{UFS_T}$  (alternativ LTFS) beschrieben werden. Diese Größe beruht auf einer Lautheits-Messung nach ITU-R BS.1770-1, welche durch die Surround- und LFE-Kanal-Gewichtung sowie die Modifizierung der K-Gewichtung durch die Kurve X an SMPTE RP200 angepasst wird. Die Sprachpassagen dieser Lautheits-Messung werden durch den Dolby Dialogue Intelligence Algorithmus ermittelt und anschließend mit einem relativen Gate bei -8 LU über den gesamten Film integriert. Dieser Wert soll maximal -30  $L_{UFS_T}$  erreichen. Die zweite Größe ist die „Theatrical Maximum Short-Term Loudness“, welche auf der gleichen angepassten Lautheitsmessung beruht, aber mit einer 3-sekündigen Fensterung über den gesamten Programminhalt ermittelt wird und einen Höchstwert von -10  $L_{UFS_T}$  nicht überschreiten soll.

Diese Modellnorm erreichte nach Simulierung an den gemessenen Filmen eine Übereinstimmung zu den Wiedergabe-Erwartungen des Autors mit einem Schwankungsbereich von 5,8 LU. Die größte positive Abweichung (zu laute Wiedergabe) betrug +2,1 LU und die größte negative Abweichung (zu leise Wiedergabe, hier bedingt durch bereits auf Level 7 zu leise klingende Filme) -3,7 LU.



Eine vergleichbare Normalisierung nach der europäischen TV-Aussteuerungs-Empfehlung EBU R 128 würde im Kino hingegen einen Schwankungsbereich von 13,0 LU und nach der TASA Trailer-Norm von 11,0 dB ergeben.

Mit dieser Pilot-Studie und der erbrachten Konzeption und Simulierung einer Modellnorm zur Aussteuerung im Kino soll eine Debatte zur Thematik angeregt werden. Es wird weiterhin empfohlen zu untersuchen, ob und in welchem definierten Maße Trailer und Werbung durch ihre kurze Programmdauer und dramaturgische Dichte von dieser Norm abweichen müssen. Die definierten Zielwerte von -30 und -10 LUFS<sub>T</sub> sollten zudem durch ausführliche Untersuchungen an Testpersonen verifiziert werden.


Die internationale Filmwirtschaft sollte sich insbesondere im Hinblick auf das digitale Kino und den damit einhergehenden schwindenden Einfluss von Kontrollinstanzen wie Dolby Laboratories auf einen Standard einigen, der die verlässliche Lautstärke-Kontinuität von Wiedergaben über 5.1 und 7.1 sowie Zukunfts-gerichtete Formate wie Dolby Atmos und Wellenfeldsynthese sichert.

Auch nach Einhaltung eines Standards wird es Filme geben, die durchweg laut sind und beim Publikum negative Reaktionen hervorrufen. Es ist dann von substanzieller Wichtigkeit zu differenzieren, ob eine Lautheits-Vorschrift falsch konzipiert wurde - oder der Film.

Aus diesem Grund sollen die letzten Worte dieser Arbeit Randy Thom gehören.

*„If you, the movie Director, don't have something interesting to say, but you're being paid several million bucks to say it, there will be an enormous temptation to yell it as loudly as you can, hoping that intensity will substitute for depth.“<sup>1</sup>*

---

<sup>1</sup>  Thom, Randy (FilmSound): Are Movies Getting Too Loud?

*„Wenn du, der Regisseur, nichts Interessantes zu sagen hast, aber mit mehreren Millionen Dollar dafür bezahlt wirst es zu sagen, dann wird die Versuchung groß sein es so laut zu brüllen wie du nur kannst, in der Hoffnungs dass Intensität Tiefe ersetzt.“*

## Literaturverzeichnis











- 📖 **Chion, Michel:** Audio-Vision (Columbia University Press, New York, 1994)
- 📖 **Chion, Michel:** Film, A Sound Art (Columbia University Press, New York, 2003)
- 📖 **Henle, Hubert:** Das Tonstudio Handbuch (Carstensen, München, 1990)
- 📖 **Zwicker & Fastl:** Psychoacoustics - Facts and Models (Springer, Berlin, 1990)
- 📄 **Allen, Ioan** (Dolby Laboratories): Are Movies too loud?  
(SMPTE Conference, Los Angeles, 1997)
- 📄 **ATSC:** A/53 Part 5 - AC3 Audio System Characteristics (Washington, 2007)
- 📄 **ATSC:** Recommended Practice A/85 - Techniques for Establishing and  
Maintaining Audio Loudness for Digital Television (Washington, 2011)
- 📄 **Benjamin, Eric:** Comparison of Objective Measures of Loudness Using Audio  
Program Material (AES Convention, Los Angeles 2002)
- 📄 **Benjamin, Eric:** Preferred Listening Levels and Acceptance Windows for Dialog  
Reproduction in the Domestic Environment  
(AES Convention, San Francisco, 2004)
- 📄 **Berliner Morgenpost:** Berlin bricht neuen Tourismus Rekord (Berlin, 15.02.2012)
- 📄 **Burg, Lothaire** (ARTE): Memo für Produzenten & Dienstleister - Lautheit von  
Fernsehprogrammen (2011)
- 📄 **DCI:** Digital Cinema System Specification (2012)
- 📄 **Dolby Laboratories:** CP750 Digital Cinema Processor Manual  
(San Francisco, 2009)
- 📄 **Dolby Laboratories:** Dolby Atmos Cinema Technical Guidelines  
(San Francisco, 2012)
- 📄 **Dolby Laboratories:** Dolby Dialogue Intelligence Reference Code User's Guide
- 📄 **Dolby Laboratories:** Technical Guidelines for Dolby Stereo Theaters  
(San Francisco, 1994)
- 📄 **DTS:** Installation and Operation Manual DTS-6D (Agoura Hills, 1998)
- 📄 **EBU:** Empfehlung R 128 - Lautheitsaussteuerung, Normalisierung und zulässiger  
Maximalpegel von Audiosignalen (Genf, 2011)
- 📄 **EBU:** TECH 3342 - Loudness Range EBU R 128 (Genf, 2011)
- 📄 **EBU:** TECH 3343 - Practical guidelines EBU R 128 (Genf, 2011)

- 📖 **ISO:** ISO Standard 21727:2004 - Cinematography - Method of measurement of perceived loudness of motion-picture audio material (2004)
- 📖 **ITU-R:** Recommendation ITU-R BS.1770-2 (Genf, 2011)
- 📖 **ITU-R:** Recommendation ITU-R BS.1770-3 (Genf, 2012)
- 📖 **Moore, Glasberg, Stone:** Why Are Commercials so Loud? - Perception and Modeling of the Loudness of Amplitude-Compressed Speech (Journal of AES, 2003)
- 📖 **Naumann, Silvio:** Entwicklung eines Verfahrens zur Generierung, Übertragung und Auswertung lautstärkerelevanter Kennwerte (Potsdam, 2008)
- 📖 **Pearsons, Bennet, Fidell:** Speech Levels in Various Noise Environments (1977)
- 📖 **Riedmiller, Lyman, Robinson** (Dolby Laboratories): Intelligent Program Loudness Measurement and Control - What Satisfies Listeners? (AES Convention, New York, 2003)
- 📖 **Skovborg & Nielsen:** Evaluation of Different Loudness Models with Music and Speech Material (AES San Francisco, 2004)
- 📖 **SMPTE:** Recommended Practice RP 200:2012 (White Plains, 2012)
- 📖 **SMPTE:** SMPTE Standard 428-2-2006 D-Cinema Distribution Master - Audio Characteristics (White Plains, 2006)
- 📖 **Sony:** SDDS Print Master Guidelines (2001)
- 📖 **Soulodre, G.A.:** Evaluation of Objective Loudness Meters (AES Convention, Berlin, 2004)
- 📖 **Soulodre, Lavoie, Norcross:** The Subjective Loudness of Typical Program Material (AES Convention, New York, 2003)
- 📖 **Spikofski** (IRT): Lautstärkemessung im Rundfunk - Stand der internationalen Standardisierung (München)
- 📖 **TASA:** The TASA Standard (1999)
- 📖 **Thom, Randy** (FilmSound): Are Movies Getting Too Loud? (<http://filmsound.org/randythom/loud-movies.htm>, zuletzt abgerufen am 25.01.2013)
- 📖 **Vinton & Robinson** (Dolby Laboratories): Automated Speech/Other Discrimination for Loudness Monitoring (AES Convention, Barcelona, 2005)
- 📖 **Ziegler, David:** Der AC-3 Standard in Theorie und Praxis (Potsdam, 2010)
- 📖 **Lerch, Christian** (Dolby Laboratories, Senior Sound Consultant, Content Services): Lautheitsmessungen bei Kinofilmen (14.11.2012)

- ☰ **Thom, Randy** (Sound-Article-List): The future of hollywood sound...  
(<http://groups.yahoo.com/group/sound-article-list/message/4650>, zuletzt abgerufen am 25.01.2013)
- ☰ **Zernicke, Jerry** (THX, Director of Professional Services): Cinema Playback Levels  
(14.12.2012)
- ☰ **BVFT**: Abhörlautstärke bei Trailermischungen  
(<http://www.bvft.de/forum/showthread.php/797-Abh%C3%B6rlautst%C3%A4rke-bei-Trailermischungen>, zuletzt abgerufen am 25.01.2013)
- ☰ **Cinefacts**: Blu-ray vs. Kino? (<http://www.cinefacts.de/Forum/210551-blu-ray-vs-kino.html>, zuletzt abgerufen am 25.01.2013)
- ☰ **DCI**: About DCI (<http://www.dcimovies.com/>, zuletzt abgerufen am 25.01.2013)
- ☰ **Dolby Laboratories**: Discovery's Deadliest Catch Customer Story  
(<http://www.dolby.com/gb/en/professional/technology/broadcast/mediameter/max-post-case-study.html>, zuletzt abgerufen am 25.01.2013)
- ☰ **DUC**: Cinema Chains - sadness at poor levels  
(<http://duc.avid.com/showthread.php?t=321916>, zuletzt abgerufen am 25.01.2013)
- ☰ **DUC**: DCP cinema volume discrepancy  
(<http://duc.avid.com/showthread.php?t=319839>, zuletzt abgerufen am 25.01.2013)
- ☰ **DUC**: DCP headroom?/levels? (<http://duc.avid.com/showthread.php?t=323775>, zuletzt abgerufen am 25.01.2013)
- ☰ **DUC**: Dolby 7. Is everybody mixing at this level?  
(<http://duc.avid.com/showthread.php?t=300904>, zuletzt abgerufen am 25.01.2013)
- ☰ **FSK**: Aufgaben und Strukturen  
(<http://www.fsk.de/index.asp?SeitID=504&TID=473>, zuletzt abgerufen am 25.01.2013)
- ☰ **Jacobs, Bruce** (Broadcast Engineering): Dialnorm - A good idea gone bad?  
(<http://broadcastengineering.com/audio/dialnorm-good-idea-gone-bad>, zuletzt abgerufen am 25.01.2013)
- ☰ **Londonnet**: Movies with parents and baby screenings currently showing in the UK (<http://www.londonnet.co.uk/films/movies-with-parents-and-baby-screenings-currently-showing-in-the-uk.html>, zuletzt abgerufen am 25.01.2013)
- ☰ **Markwick, Kevin** (Cinema Owner Confidential): Can you hear me at the back?  
(<http://kevinmarkwick.co.uk/2011/08/09/can-you-hear-me-at-the-back/>, zuletzt abgerufen am 25.01.2013)

- 📄 **Media Channel:** Sendungen in 5.1 Dolby Digital Surround Sound - TV und Hörfunk (<http://dolby-digital.the-media-channel.com/>, zuletzt abgerufen am 25.01.2013)
- 📄 **Moulton, Dave** (TVTechnology): LKFS & The CALM Act (<http://www.tvtechnology.com/inside-audio/0176/lkfs-amp-the-calm-act/208806>, zuletzt abgerufen am 25.01.2013)
- 📄 **TASA:** TASA History (<http://www.tasatrailers.org/history.html>, zuletzt abgerufen am 25.01.2013)
- 📄 **TASA:** Who are we? (<http://www.tasatrailers.org/whoarewe.html>, zuletzt abgerufen am 25.01.2013)

# Abbildungsverzeichnis

<b>Abbildung 1: Komfort-Zone TV Zuschauer</b> .....	10
(aus  Riedmiller, Lyman, Robinson (Dolby Laboratories) S.6)	
<b>Abbildung 2: Übereinstimmung bei Lautstärke-Wahrnehmung von Sprache und Geräuschen</b> .....	11
(aus  Riedmiller, Lyman, Robinson (Dolby Laboratories) S.6)	
<b>Abbildung 3: Gewichtungs-Kurven</b> .....	29
(aus  Skovenborg & Nielsen: Evaluation of Different Loudness Models with Music and Speech Material, S. 7)	
<b>Abbildung 4: Dialog-Normalisierung versch. Programme</b> .....	34
(Quelle: Dolby Laboratories)	
<b>Abbildung 5: Blockschaltbild Dialogue Intelligence Algorithmus</b> .....	35
(aus  Vinton & Robinson (Dolby Laboratories): Automated Speech/Other Discrimination, S.2)	
<b>Abbildung 6: Lautsprecheranordnung ITU-R BS.775</b> .....	41
(aus  ITU-R: Recommendation ITU-R BS.1770-3, S. 12)	
<b>Abbildung 7: Korrelation subj. Wahrnehmung zu ITU-R BS.1770-1</b> .....	41
(aus  ITU-R: Recommendation ITU-R BS.1770-3, S. 16)	
<b>Abbildung 8: K-Gewichtung</b> .....	42
(aus  EBU: TECH 3343 - Practical guidelines EBU R 128, S. 9)	
<b>Abbildung 9: Ermittlung des True Peak Level</b> .....	44
(aus  ITU-R: Recommendation ITU-R BS.1770-2, S. 18)	
<b>Abbildung 10: Schematische Darstellung von Peak- und Lautheits-Normalisation</b> .....	46
(aus  EBU: TECH 3343 - Practical guidelines EBU R 128, S. 16)	
<b>Abbildung 11: Differenz von Programmlautheit und Dialoglevel nach BS.1770-1 (LU)</b> .....	60
<b>Abbildung 12: Entwicklung der Lautheit nach BS.1770-1 im Film „Alle Anderen“</b> .....	61
(erstellt durch Dolby Media Meter 2)	
<b>Abbildung 13: Entwicklung der Lautheit nach BS.1770-1 im Film „The International“</b> .....	62
(erstellt durch Dolby Media Meter 2)	
<b>Abbildung 14: Entwicklung der Lautheit nach BS.1770-1 im Film „Lissi und der wilde Kaiser“</b> .....	63
(erstellt durch Dolby Media Meter 2)	
<b>Abbildung 15: Short-Term Loudness und True Peak Level im Film „Lissi und der wilde Kaiser“</b> ..	63
(erstellt durch NUGEN LMB)	
<b>Abbildung 16: Spracherkennung von Rap-Passagen im Film „Blutzbrüdaz“</b> .....	64
(erstellt durch Dolby Media Meter 2)	
<b>Abbildung 17: Implementierung Dolby Dialogue Intelligence</b> .....	65
(aus  Dolby Laboratories: Dolby Dialogue Intelligence Reference Code User’s Guide, S. 17, Layout geändert)	
<b>Abbildung 18: Subjektive Pegelung Autor (Dolby Skala)</b> .....	67
<b>Abbildung 19: Entwicklung des Dialoglevels nach BS.1770-1 im Film „R U There“</b> .....	73
(erstellt durch Dolby Media Meter 2)	
<b>Abbildung 20: Frequenzanalyse des fallenden Kronleuchters in „The International“</b> .....	85
(erstellt durch Waves PAZ Frequency )	
<b>Abbildung 21: Finale Abweichungen von der Ideallautstärke vor und nach Modellnorm</b> .....	87

# Tabellenverzeichnis

<b>Tabelle 1: DI Features und deren Blockgrößen</b> .....	36
(aus [1] Vinton & Robinson (Dolby Laboratories): Automated Speech/Other Discrimination, S. 3)	
<b>Tabelle 2: Filmdatenbank, Allgemeine Informationen</b> .....	55
(Zusatzinformationen aus [2] Internet Movie Database: www.imdb.com)	
<b>Tabelle 3: Filmdatenbank, Messungsergebnisse</b> .....	56
<b>Tabelle 4: Nach subj. Pegelung resultierende Programmlautheiten und Dialoglevel</b> .....	68
<b>Tabelle 5: Abweichungen vom Mittelwert und resultierender Schwankungsbereich</b> .....	69
<b>Tabelle 6: Auswirkung des relativen Gates bei -8 LU</b> .....	75
<b>Tabelle 7: Short-Term und Momentary Loudness besonders lauter Filmszenen</b> .....	78
<b>Tabelle 8: Abweichungen von der Modellnorm</b> .....	86





## **Selbstständigkeitserklärung**

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe. Insbesondere versichere ich, dass ich alle wörtlichen und sinngemäßen Übernahmen aus anderen Werken als solche kenntlich gemacht habe.

---