

# **Heath Ledger in „The Dark Knight“**

## **Ein analytischer Vergleich zwischen Originalstimme und Synchronisation in Bezug auf Hörerurteile**

Masterarbeit

im Fach

Speech Science  
Spezialisierung Sprechwissenschaft

dem Fachbereich  
Germanistik und Kunstwissenschaften  
der Philipps-Universität Marburg

vorgelegt von

Christian Haas

aus Göppingen

Marburg, 2009

Veröffentlicht auf  
[www.dgss.de](http://www.dgss.de)  
dem Internetauftritt der Deutschen Gesellschaft für  
Sprechwissenschaft und Sprecherziehung (DGSS) e.V.

## Danksagung

Diese Arbeit wäre in der vorliegenden Form nicht ohne die Menschen möglich gewesen, die mir dafür in unterschiedlichster Form ihr Wissen, ihre Zeit und ihre Fähigkeiten zur Verfügung stellten.

Zuallererst gilt mein Dank Prof. Dr. Christa M. Heilmann, welche die Arbeit über den gesamten Entstehungsprozess betreut und durch zahlreiche wertvolle Hinweise und kritische Anmerkungen wesentlich dazu beigetragen hat, dass aus einer spontanen Idee meinerseits eine konkrete Arbeit werden konnte. Ebenso danke ich ihr für ihre engagierte Lehre, nicht nur in den letzten zwei Jahren. Für die Übernahme des Zweitgutachtens danke ich Prof. Dr. Richard Wiese, der auch als Lehrender neben Prof. Dr. Angelika Braun, Prof. Dr. Joachim Herrgen und Dr. Luitgard Soni wesentlich zu meiner akademischen Ausbildung beigetragen hat.

Ganz besonders danke ich meinen beiden Kommilitoninnen Marie-Theres Braun und Svenja Völkert für zwei sehr schöne Masterjahre und den regen fachlichen Austausch, der verhindert hat, dass ich im Arbeitsdschungel den Überblick verliere. Ein spezieller Dank geht an dieser Stelle auch an Annika Johannsen, die durch eine kleine Bemerkung die erste Saat zur Idee dieser Arbeit legte.

Meinen Kontrollhörerinnen Friederike Beckmann und Carolin Zingel danke ich für ihr Engagement, ihre wertvollen Hinweise und ihr „offenes Ohr“ in der auditiven Analyse. Dass ich die akustische Analyse ohne große Schwierigkeiten in Angriff nehmen konnte, verdanke ich Hanni Schnell und ihrer kompetenten Einführung in die Funktionsweise der nötigen Programme. Außerdem danke ich Monika Wiebe für die Beantwortung meiner statistischen Fragen, sowie David Breier für seine Hinweise zu Layout und Textverarbeitung.

Conny Krause und Alexander Dröge haben durch ihre zahlreichen Anmerkungen und die wertvolle Kritik die Korrekturphase der Masterarbeit entscheidend vorangebracht. Herzlichen Dank dafür.

Daneben danke ich all denjenigen, die durch ihre Freundschaft, ihre Motivation und etwas Ablenkung zur rechten Zeit so manchen Geistesblitz erst ermöglicht haben. Wer von euch oben nicht schon erwähnt ist, darf sich hiermit angesprochen fühlen.

Mein größter Dank aber gilt abschließend meinen Eltern sowie meiner Großmutter für ihre liebevolle Begleitung und Unterstützung in allen Lebenslagen, ohne die mein Studium nicht möglich gewesen wäre. Ihnen ist diese Arbeit gewidmet.

## Inhaltsverzeichnis

<b>1. Einleitung: Die Synchronisation von Heath Ledger in „The Dark Knight“ .....</b>	<b>6</b>
<b>2. Theoretische Grundlagen.....</b>	<b>9</b>
2.1 Die menschliche Sprechstimme.....	9
2.2 Der Sprechausdruck.....	10
2.2.1 Sprechausdruck und Höreindruck.....	10
2.2.2 Die Parameter des Sprechausdrucks.....	13
2.3 Sprechausdruck und Persönlichkeitsbeurteilung – bisherige Untersuchungen.....	20
2.4 Filmsynchronisation in Deutschland.....	22
<b>3. Ein Vergleich zwischen Originalstimme und Synchronisation.....</b>	<b>27</b>
3.1 Sprechausdruck im synchronisierten Film – Vorüberlegung, Hypothesen.....	27
3.2 Material und Methoden.....	28
3.2.1 Die Vergleichsgrundlage: Heath Ledger als Joker in „The Dark Knight“.....	28
3.2.2 Auswahl, Gewinnung und Verteilung der Hörerurteile.....	31
3.2.3 Auswahl und Gewinnung des Sprachmaterials.....	36
3.2.4 Analyse des Sprachmaterials.....	38
3.2.4.1 Auditive Analyse.....	38
3.2.4.2 Akustische Analyse.....	40
3.2.4.3 Analyse der temporalen Parameter.....	41
3.3 Ergebnisse.....	43
3.3.1 Video-Szene.....	44
3.3.2 Party-Szene.....	46
3.3.3 Vater-Szene.....	49
3.3.4 Gordon-Szene.....	50
3.3.5 Batman-Szene.....	53
3.3.6 Hospital-Szene.....	55
3.3.7 Zusammenfassung.....	56
<b>4. Diskussion.....</b>	<b>59</b>
4.1 Ergebnisdiskussion im Hinblick auf bisherige Ergebnisse.....	59
4.2 Schlussfolgerungen und Ausblick.....	65

<b>5. Literaturverzeichnis.....</b>	<b>69</b>
<b>Anhang.....</b>	<b>73</b>
A. Für die Analyse verwendete Userkommentare auf YouTube.com.....	74
B. Daten der Akustischen Analyse.....	82

## Abbildungsverzeichnis

<b>Abbildung</b>	<b>Beschriftung</b>	<b>Seite</b>
Abbildung 3.1	Anzahl der Kommentare in Relation zum Abgabezeitraum	33
Abbildung 3.2	Grundfrequenzverlauf des Satzes: <i>You see, this is how crazy batman's made Gotham</i>	45
Abbildung 3.3	Grundfrequenzverlauf des Satzes: <i>Seht ihr, so verrückt hat Batman Gotham gemacht</i>	45
Abbildung 3.4	Grundfrequenzverlauf des Satzes: <i>Good evening Ladies and Gentlemen</i>	47
Abbildung 3.5	Grundfrequenzverlauf des Satzes: <i>Guten Abend, Ladys und Gentlemen</i>	48
Abbildung 3.6	F <sub>0</sub> -Verlauf von <i>depending on the time, he may be on one spot or several</i>	51
Abbildung 3.7	F <sub>0</sub> -Verlauf von <i>je nachdem ist er entweder an einem Ort oder gleich an mehreren</i>	52

Es ist keine Sprache noch Rede, da man nicht ihre Stimme höre.  
*Die Bibel, Psalm 19,4*

## 1. Einleitung:

### Die Synchronisation von Heath Ledger in „The Dark Knight“

„Die vielbeschworene Einheit von Körper und Stimme wurde mir auf verwirrende Weise erschüttert, als ich zum ersten Mal [...] John Wayne mit seiner eigenen geschmeidigen Stimme reden hörte, statt der eher abstoßenden rauhen Synchronstimme seines deutschen Sprechers Arnold Marquis, die seinen Rollen und seiner Erscheinung soviel mehr entsprach.“ (Frieda Grafe, zitiert nach Maier 1997: 13)

Lange Zeit waren Eindruckserlebnisse wie das Frieda Grafes in Deutschland die Ausnahme. Denn die Synchronisation ausländischer Spielfilme ist ein so populäres Mittel zur Filmübertragung geworden, dass sie von den Zuschauern in der Regel kaum noch hinterfragt wird (vgl. Bräutigam 2001: 6). Erst seit das Medium DVD den Vergleich der unterschiedlichen Tonspuren ermöglicht, rückt stärker ins Bewusstsein, dass die „Einheit von Körper und Stimme“ im synchronisierten Film eine Illusion darstellt (vgl. Blaseio 2003: 160f.). Es fällt vielmehr auf, dass gar nicht wenige ausländische Schauspieler Synchronstimmen besitzen, die sich von ihrer eigenen bemerkenswert unterscheiden. Interessant ist dabei die Beobachtung, dass diese Stimmen nicht nur eine Übersetzung des Inhalts vermitteln, sondern durch ihre klanglichen Eigenschaften auch die Wirkung der jeweiligen Rolle mit beeinflussen (vgl. Bräutigam 2001: 29; Eckert/Laver 1994: 140, Blaseio 2003: 163 u. 169). Wie viel stärker diese Tatsache durch den Zuschauer heutzutage wahrgenommen wird, zeigt sich vor allem auch darin, dass Stimmenvergleiche zwischen Original und Synchronisation in einschlägigen Foren und Kommentarseiten des Internets immer wieder Gegenstand von Diskussionen sind.

So auch im Fall des Films „The Dark Knight“ des Regisseurs Christopher Nolan. Schon vor dem offiziellen deutschen Kinostart am 21. August 2008 gaben Internetnutzer unter anderem auch ihre Meinung zu der durch die vorab veröffentlichten Trailer erwartbaren Synchronisation ab. Insbesondere die Interpretation der Figur des Joker durch den australischen Schauspieler Heath Ledger wurde mit Spannung erwartet und auch der Eindruck seiner deutschen Stimme von den Fans bewertet. Diese Diskussion vererbte auf einschlägigen Seiten auch nach Start des Films nicht und dauert bis heute an. So lud im Juni 2008 ein Internetnutzer im Videoportal „YouTube“ ein Video hoch, in dem er der Öffentlichkeit eine Szene des Films zugänglich machte und die deutsche und

englische Sprachfassung so schnitt, dass ein direkter Stimmenvergleich zu einzelnen Dialogzeilen des „Jokers“ möglich wurde. Daraufhin gaben zahlreiche User nach Betrachtung des Videos über die Kommentarfunktion der Seite ihren unmittelbaren<sup>1</sup> Eindruck der Stimmen schriftlich wieder<sup>2</sup>. Dabei fällt auf, dass einige User in der Diskussion um die stimmliche Leistung Heath Ledgers und Simon Jägers, der Synchronstimme, auch eine Beurteilung des „Klangeindrucks“ der Stimme mitliefern, ähnlich wie Frieda Grafe im obigen Beispiel. So schreibt etwa ein Nutzer: „ich finde beide gleichgut ...aber auf deutsch kommt der joker iwie noch verrückter vor und das passt finde ich mehr“<sup>3</sup>.

Obwohl der Eindruck und die Wirkung der unterschiedlichen Stimmen also durchaus in der Öffentlichkeit ein Thema ist, hat sich die Wissenschaft bisher kaum mit diesem Phänomen der Synchronisation auseinandergesetzt. Zwar gibt es einige mehr oder weniger wissenschaftliche Arbeiten zum Thema Synchronisation an sich, diese beschäftigen sich aber eher mit den inhaltlichen Veränderungen, die eine solche sprachliche Bearbeitung nach sich zieht, zum Beispiel durch Vergleich der Dialogzeilen (vgl. Hesse-Quack 1961, Maier 1997, Pruys 1997, sowie die Anmerkungen von Blaseio 2003: 161). Die unterschiedlichen Konsequenzen, die der *Stimmaustausch* möglicherweise auf die Wirkung eines Films hat, werden wenn nur am Rande thematisiert (vgl. Maier 1997: 107f.). Dabei fällt bei genauerer Betrachtung der Kommentare auf YouTube.com auf, dass diese vor allem in Bezug auf die deutsche Stimme erstaunlich übereinstimmend sind (vgl. Kapitel 3.2.2).

Haben diese gemeinsamen Wahrnehmungen deshalb etwa ihre Grundlage tatsächlich in den verschiedenen Stimmen? Und wodurch kommt dies möglicherweise zustande?

Mit diesen Fragen beschäftigt sich immer wieder die Disziplin der Sprechwirkungsforschung (vgl. etwa Stock 1991, Eckert/Laver 1994: 158). Bisher geschah dies jedoch überwiegend anhand vorgegebener Beurteilungskategorien oder Fragebögen. Die Eindrucksbildung war somit mehr oder weniger geleitet (vgl. Stock 1991: 53) oder durch zu große Reflektierung über das eigene Verhalten verfälscht (vgl. Franz 2008: 9). Dies lag vor allem daran, dass „spontane“ Eindrücke aus Felduntersuchungen, die nicht im Rahmen eines Experiments abgegeben wurden,

---

<sup>1</sup> „Unmittelbar“ deshalb, weil davon auszugehen ist, dass Nutzer die Kommentarfunktion direkt nach Betrachten des Videos nutzen oder zumindest in recht kurzem zeitlichem Anschluss daran.

<sup>2</sup> vgl. auch im Folgenden die Kommentare auf der Seite „YouTube – The Dark Knight Synchro-Vergleich (Englisch/Deutsch)“ unter <http://www.youtube.com/watch?gl=DE&hl=de&v=kW3KP3NgIBs> (21.07.09). Kommentare, auf die in der Arbeit direkt Bezug genommen wird, sind im Anhang aufgelistet.

<sup>3</sup> Kommentar von koenigseggfan3.

schwer zu bekommen waren (vgl. Stock 1991: 21). Mit den obigen Internetkommentaren schien nun jedoch auf recht einfachem Wege ein solches spontanes und unmittelbares Urteilskorpus gefunden (vgl. Franz 2008: 9). Hier geben Hörer zum ersten Mal in einer Nicht-Laborsituation unterschiedliche Urteile über die Wirkung zweier Stimmen ab, die auch noch recht konstant zu sein scheinen.

Daraus ergeben sich weiter folgende Fragen: Sind die beschriebenen stimmlichen Eindrücke völlig willkürlich? Oder lassen sich diese Wirkungen durch eine vergleichende Analyse beider Stimmen aus dem Sprachsignal heraus rekonstruieren und nachvollziehbar machen? Und lässt sich dann erklären, welche Elemente des Sprechausdrucks die unterschiedlichen Wirkungen hervorrufen, trotz vieler Variablen? Dem will die vorliegende Arbeit nachgehen.

Vorangestellt wird der Analyse in Kapitel 2 ein theoretischer Teil, der die grundlegenden Erkenntnisse der Sprechausdrucks- und Sprechwirkungsforschung erläutert. Ebenso wird hier auf die für die Fragestellung der Arbeit wichtigen Hintergründe der Filmsynchronisation eingegangen. Anschließend werden in Kapitel 3 – nach einer methodischen Einführung – die Stimmen Heath Ledgers und Simon Jägers mittels einer auditiven und akustischen Analyse verglichen und auffällige Unterschiede in der Ausprägung einzelner Merkmale herausgearbeitet. Den Abschluss der Arbeit bildet in Kapitel 4 eine ausführliche Methoden- und Ergebnisdiskussion, die versucht, die Ergebnisse aus Kapitel 3 zu Ergebnissen der Sprechwirkungsforschung in Beziehung zu setzen und Erklärungsansätze für das Zustandekommen der weitgehend übereinstimmenden Nutzerurteile herauszuarbeiten.



## 2. Theoretische Grundlagen

### 2.1 Die menschliche Sprechstimme

Um zu verstehen, wie verschiedene Klangqualitäten in der menschlichen Stimme zum Ausdruck gebracht werden können, sollen im Folgenden die organischen Grundlagen für das Zustandekommen der Sprechstimme des Menschen kurz skizziert werden. Die Bezeichnung „menschliche Sprechstimme“ für die „Stimme, mit der Menschen sprechen“ wird hier verwendet, um sie sowohl von tierischen Lautäußerungen (vgl. Fischer/Hartlieb 1984: 35) als auch von der Singstimme des Menschen abzugrenzen (vgl. Zilliken 1990: 8f.).

Organische Grundlage für alle stimmlichen Äußerungen des Menschen sind die Lunge, der Kehlkopf und der Bereich oberhalb des Kehlkopfes bis zu den Lippen, das sogenannte Ansatzrohr. Luftdruckdifferenz und Luftströmung aus der Lunge liefern die notwendige Energie zur Erzeugung des Sprachschalls, die im Kehlkopf stattfindet. Beim Kehlkopf handelt es sich um ein „mehrfach geteiltes Knorpelgerüst“ (Fischer/Hartlieb 1984: 35), das sich am oberen Ende der Luftröhre befindet. Die einzelnen Teile dieses Knorpelgerüsts sind so durch Muskeln fixiert, dass sie sich in verschiedene Richtungen zueinander bewegen lassen und so insgesamt die Spannungsverhältnisse im Kehlkopf verändern können (vgl. Pétursson/Neppert 2002: 63ff.). Die Basis hierfür bildet der sogenannte Ringknorpel. Auf seinem vorderen Rand befindet sich der Schildknorpel, dessen beide Platten als „Adamsapfel“ sicht- bzw. spürbar sind. Hinten auf dem Ringknorpel sitzen links und rechts die beiden Aryknorpel. Diese werden auch Stellknorpel genannt, da sie sich zueinander in mehrere Richtungen verstellen lassen. Zwischen den beiden Aryknorpeln und der Innenseite des Schildknorpels sind die beiden wichtigsten Muskelzüge des Kehlkopfes gespannt, die sogenannten Stimmlippen. Liegen diese durch Stellung der Aryknorpel aneinander, können sie durch die Luftströmung aus der Lunge in Schwingung versetzt werden und der Sprachschall entsteht. Dieser Vorgang wird als *Phonation* bezeichnet<sup>4</sup>. Der entstandene Schall kann in seiner Frequenz und Ausprägung durch die Aktivität der einzelnen Muskeln und Stellung der Knorpel unterschiedlich verändert werden. So bewirkt etwa ein Kippen des Schildknorpels nach vorn eine höhere Spannung der

---

<sup>4</sup> Eine ausführliche Erläuterung der (vermuteten) physikalischen Vorgänge, die zur Schwingung der Stimmlippen führen, findet sich z.B. bei Eckert (2004: 21f.), Pétursson und Neppert (2002: 70ff.) und Laver (1980: 95f.).

Stimmlippen, was deren Schwingungszahl verändert und den Ton höher erscheinen lässt<sup>5</sup> (vgl. Pétursson/Neppert 2002: 65f.). Außerdem spielen die Veränderung der Lage dieser Organe und unterschiedliche Spannungsverhältnisse der beteiligten Muskeln auch eine große Rolle bei der Entstehung verschiedener Stimmqualitäten, worauf später noch ausführlicher eingegangen wird. Wichtig ist für den Moment nur die Unterscheidung zwischen der Stellung in Ruheatmung, bei welcher der Spalt zwischen den Stimmlippen, die sogenannte Glottis, geöffnet ist, und der Phonationsstellung. Hier ist die Glottis geschlossen, die Stimmlippen liegen nebeneinander und können in Schwingung versetzt werden.

Der so entstandene Primärton versetzt nun seinerseits das Ansatzrohr, die „Gesamtheit der Hohlräume oberhalb des Kehlkopfes“ (Pétursson/Neppert 2002: 76), in Schwingung und erhält so die resonante Ausprägung, die vom menschlichen Ohr wahrgenommen wird. Rachen-, Mund- und Nasenraum fungieren damit als Resonatoren des Sprachschalls. Schließlich wird der aus dem Kehlkopf kommende Luftstrom (und damit der Schall) durch die Artikulationsorgane Zunge, Lippen, Zähne und Unterkiefer so geformt, dass einzelne Laute entstehen und der Sprechschall hörbar wird.

## 2.2 Der Sprechausdruck

### 2.2.1 Sprechausdruck und Höreindruck

Wenn Menschen miteinander sprechen, so äußern sie wie oben beschrieben nicht einfach nur Laute und formen diese zu Wörtern und Sätzen. Auch durch die Art des Sprechens drückt der Sprecher gegenüber dem Hörer etwas aus. Oder wie Geißner (1984b: 10) es formuliert: „Im Sprechen ist immer zugleich Organisches (Sprechwerkzeuge/Hörorgan), immer Sinnliches (Ausdruckshaftes) und immer Sprachliches (Bedeutungshaftes)“. Dieses „Sinnliche“ bedeutet, dass das Sprechen nie vom Sprechenden selbst zu trennen ist. Es wird immer „persongebunden [...] leibhaft vollzogen“ (Geißner 1988: 116), das heißt, der Sprecher gibt immer etwas von sich, seiner Persönlichkeit in die Kommunikation, in das Sprechen mit hinein. Fischer und Hartlieb (1984: 42) beschreiben dies so: „Von seiner Person, vielleicht von der „Persönlichkeit“ akzentuiert, gehen Wirkungen aus, die als spezifische Mittel sich in Stimmklang und Artikulation [...] manifestieren“. Da Sprechen immer mit einem

---

<sup>5</sup> Auf den Zusammenhang zwischen Tonhöhe und Frequenz werde ich an anderer Stelle genauer eingehen (vgl. Kapitel 3.2.4.2).

Sprecher verbunden geschieht, werden darüber auch unter anderem dessen momentane Einstellungen und Gefühle hörbar, da sich das körpereigene, psychische Empfinden auch oder gerade auf den sehr feinnervigen Ablauf der Muskelbewegungen beim Sprechen auswirkt<sup>6</sup>. „Im WIE der Stimme kommt die jeweilige Stimmung unwillkürlich zum Ausdruck“ (Geißner 1988: 126). Damit erhalten nicht nur die artikulierte Laute im Sprechen ihre festgelegte Ausdrucksfunktion, sondern alle stimmlichen Merkmale übermitteln Informationen an den Hörer und werden so zum Ausdrucksträger des Sprechers selbst, abhängig von seiner Konstitution und der Situation, in der sich der Sprecher befindet (vgl. Geißner 1984b: 16). Damit folgt die vorliegende Arbeit der (Sprech)Ausdrucksdefinition von Geißner:

„Wenn von ‚Ausdruck‘ gesprochen wird, so ist damit keine besondere Emphase, keine exaltierte Expressivität gemeint, sondern jene aus dem leibhaften Vollzug kommende ‚Prägung‘, die in jeder mündlichen Äußerung unmittelbar und unlösbar geäußert wird, und die sowohl den semantisierbaren als auch den nicht-semantisierbaren Sinn der Äußerung mitbestimmt“ (Geißner 1988: 117).

Den Ausdruck bilden hierbei die vom Hörer wahrnehmbaren Veränderungen der stimmlichen und artikulatorischen Vorgänge vom Kehlkopf bis zum Ansatzrohr<sup>7</sup>, die sich als „verschiedene Realisationsweisen im Sprechschall“ niederschlagen (Bose 2003: 32). Diesen Realisationsweisen kann der Hörer verschiedene „Bedeutungen“ – i.S.v. Gedanken, Gefühlen und Wünschen des Sprechers (vgl. Geißner 1988: 118) – zuordnen<sup>8</sup>, weil sie innerhalb einer Sprachgemeinschaft als Sprech- und Hörmuster konventionalisiert sind. Das heißt, es existieren überindividuelle Sprechausdrucksmuster, die erlernt und damit als Höreindruck erkennbar sind (vgl. Geißner 1988: 120, Bose 2003: 91). Gerade bei Emotionen wurde und wird diese „Mustererkennung“ sehr ausführlich untersucht<sup>9</sup>, aber auch die Übermittlung von Persönlichkeitseigenschaften wird als Bestandteil des Sprechausdrucks angesehen<sup>10</sup>.

<sup>6</sup> Vgl. hierzu Zilliken (1991: 10) und Eckert/Laver (1994: 6).

<sup>7</sup> Der Stimmklang macht hierbei nur einen Teil des Komplexes aus. Gerade dadurch, dass sämtliche auditiven Eigenschaften der Stimme vom Hörer bestimmten Mustern zugeordnet werden, sind sie insgesamt „Ausdruck“ von etwas, bekommen also Zeichencharakter. Auch außersprachliche Mittel entfalten so eine bestimmte Wirkung auf den Hörer (vgl. etwa Stock 1991: 22f.). Sprecher und Hörer schöpfen dabei aus einem gemeinsamen Zeicheninventar (vgl. Geißner 1984: 10f.).

<sup>8</sup> Vgl. zu diesem Aspekt auch Schulz von Thun (1981), der einer Nachricht des Sprechers neben dem Sachinhalt noch drei weitere Ebenen zuordnet. So spricht er u.a. von einer Selbstoffenbarungsseite der Nachricht und einer Beziehungsseite. Beides sind zusätzliche Informationen, die der Hörer aus der Art des Wortlauts/Stimmklangs ziehen kann.

<sup>9</sup> Vgl. z.B. Braun/Heilmann (2006), Kienast (2002), Kranich (2002), Kehrein (2002) und Scherer (1982) sowie Stock (1991: 34).

<sup>10</sup> Vgl. Kapitel 2.3 dieser Arbeit sowie Lowens Beispiel in Zilliken (1991: 11) und die entsprechenden Anmerkungen.

Obwohl dieser Ausdruck im dynamischen, melodischen, artikulatorischen und temporalen WIE des Sprechens auch eine linguistische Funktion hat<sup>11</sup> (vgl. Geißner 1984b: 13), geht er doch weiter als diese. Denn der Begriff Sprechausdruck umfasst ebenso die personenbezogenen Merkmale des Sprechens in ihrer sozial-expressiven Funktion und Wirkung (vgl. Heilmann 2002: 20), abhängig von Situation und Rolle des Sprechers (vgl. Geißner 1988: 121). Sprechausdruck hat somit auch eine personal-kommunikativ-intentionale Funktion (vgl. auch Zilliken 1991: 12, Stock/Suttner 1991: 66, Bose 2003: 38). Da dieser kommunikativ-psychosoziale Wirkungsaspekt in der vorliegenden Arbeit im Vordergrund steht, wird hier für diesen Funktionskomplex aus sprecherischen Merkmalen der Begriff *Sprechausdruck* verwendet, auch als Abgrenzung zu rein linguistischen Termini wie Prosodie und Intonation<sup>12</sup>. Außerdem ist der Begriff in der Sprechwissenschaft „verbreitet und akzeptiert, und es liegen unter dieser Bezeichnung Angaben über die Merkmalsbereiche sprecherischen Gestaltens [...] vor“ (Bose 2003: 32), deren Untersuchung auch diese Arbeit zum Gegenstand hat.

Neben der Beschaffenheit des Sprechausdrucks – dessen genaue Parameter im nächsten Kapitel erläutert werden – spielt für die hier durchgeführte Untersuchung vor allem sein Wirkungsaspekt eine Rolle. Das heißt, es geht um die Frage, wie neben dem Ausdruck in der Kommunikation ein *Höreindruck* beim Hörer entsteht, der diesen unter anderem zu einer bestimmten Bewertung des Stimmklangs bewegt, mit der Aussage, die Person spräche in einer bestimmten Art und Weise<sup>13</sup>.

Die Beobachtung von Alltagskommunikation hat gezeigt, dass ein solcher Höreindruck ebenfalls ein Komplex aus allen hör- und sichtbaren Merkmalen des Sprechers ist<sup>14</sup> (vgl. U. Geißner 1984: 128f.; Stock 1991: 22f.). Dies bedeutet jedoch nicht, dass dieser Eindruck nur aus dem Sammeln einzelner Daten, ein Höreindruck also aus der Zusammensetzung einzelner stimmlicher Parameter entstünde. Ursula Geißner bemerkt dazu: „Der Prozeß, der zu Eindrücken führt, wird besser von seiner funktionalen Zielsetzung deutlich: Menschen beobachten sich weder objektiv – sondern immer schon interpretierend – noch ziellos“ (1984: 131). Die „funktionale Zielsetzung“ für das

<sup>11</sup> Indem er zum Beispiel einzelne Teile einer Äußerung durch Veränderung des Melodieverlaufs gliedert (vgl. z.B. Geißner 1988: 119, Pétursson/Neppert 2002: 152).

<sup>12</sup> Auch Niebuhr (vgl. 2007: 13) merkt an, dass die „Sprechmelodie“ für sich allein oft anders interpretiert wird als im situativen Kontext (vgl. zu dieser terminologischen Unterscheidung auch Heilmann (2002: 21), sowie die ausführliche Begriffsdiskussion in Bose (2003: 28ff.)).

<sup>13</sup> In unserem Beispiel etwa, dass die Synchronstimme des Jokers „wahnsinniger“ klinge, d.h. auf den Hörer aufgrund bestimmter sprecherischer Merkmale einen klanglichen Eindruck bewirkt, welcher die Assoziation „wahnsinnig“ mehr hervorruft als die Vergleichsstimme.

<sup>14</sup> Vgl. zur Gleichzeitigkeit von Sprechausdruck und Höreindruck auch Geißner (1988: 124).

Sammeln von Eindrücken ist die Möglichkeit zur Voraussage von Handlungen des anderen. Zu diesem Zweck wollen Menschen andere einschätzen und beziehen dabei nicht nur das Sicht- und Hörbare mit ein. Ein Eindruck bildet sich immer auch aus den persönlichen Erfahrungen, Projektionen und Übertragungen des Beurteilenden (vgl. U. Geißner 1984: 133). Dies gilt auch für die Eindrucksbildung aufgrund von sprecherischem Ausdruck. Das „Bild“, das man sich von einer Person vor seinem geistigen Auge aufgrund anderer Faktoren und Erfahrungen bereits gemacht hat, wird mit dem auditiven Eindruck ihrer Stimme gleichgesetzt und dieser durch dieses Bild interpretiert (vgl. Stock 1991: 52). Denn Menschen sind nicht bloße Sender und Empfänger der Sprache, sondern kommunizieren als sozial-historische Wesen (vgl. Stock 1991: 27f.). Alle scheinbar deskriptiven Aussagen über Stimme sind deshalb mit anderen (z.B. über Verhalten) verbunden (vgl. U. Geißner 1984: 131), es bilden sich für eine Sprachgemeinschaft sogenannte Hörmuster heraus. Diese Komplexität mehrerer Faktoren und Muster muss bei der Beurteilung sowohl des Zustandekommens eines Eindrucks als auch von dessen Wirkung mit berücksichtigt werden.

Dies macht zwei Konsequenzen für die hier durchgeführte Analyse deutlich: Zum einen darf bei der Untersuchung der Wirkung eines Sprechausdrucks nie nur dieser alleine in eine Interpretation darüber mit einbezogen werden<sup>15</sup>. Nichtsdestotrotz ist es zweitens dennoch sinnvoll, den Sprechausdruck als Teil des Höreindrucks zu untersuchen, um dieses Komplexgebilde anschließend bis zu einem gewissen Grad nachvollziehbar machen zu können. Dazu muss aber der Sprechausdruck selbst zunächst in Parameter zerlegt werden, welche die zugrundeliegenden sprecherischen Vorgänge verdeutlichen. Erst dann lässt sich rückführend seine Zusammensetzung als auditiv wahrnehmbares Gebilde jenseits von schwammigen Alltagsmetaphern beschreiben (vgl. U. Geißner 1984: 131f.; Stock 1991: 24).

### 2.2.2 Die Parameter des Sprechausdrucks

Die Frage ist nun, wie sich dieser Sprechausdruck, der als Komplex beim Hörer ankommt, so beschreiben lässt, dass die einzelnen Vorgänge, die ihn konstituieren, messbar und damit nachvollziehbar gemacht werden können (vgl. Brown 1982: 211).

Dazu muss der Sprechschall zunächst in seine auditiv wahrnehmbaren und akustisch messbaren Einzelkomponenten zerlegt werden. Eine erste Klassifikation dieser

---

<sup>15</sup> Hierauf wird in Kapitel 4 bei der Interpretation der Ergebnisse näher eingegangen.

einzelnen Parameter stammt von dem Sprecherzieher Erich Drach (vgl. Fährmann 1960: 4f.) und wurde in der Sprechwissenschaft in den folgenden Jahren systematisch weiterentwickelt (vgl. Bose 2003: 38). Geißner etwa fasst die einzelnen Parameter des Sprechausdrucks in seinem „Ausdrucks katalog“ zusammen. (vgl. z.B. 1989: 76f.). Als „Grundqualitäten“ bezieht er dabei die Elemente Melodik, Dynamik (Lautstärke), Tempo und Artikulation mit ein. Andere Autoren nehmen hier noch die Merkmale Sprechspannung (vgl. z.B. Heilmann 2002, Bose 2003) und Stimmqualität bzw. Stimmklang (vgl. Laver 1980, Eckert/Laver 1994, Kranich 2002) mit hinzu. Worin manifestieren sich nun diese Qualitäten und wie wirken sie zum Sprechausdruck zusammen?

Im konkreten Sprechereignis gibt es diese Parameter zwar nicht einzeln – und sie bedingen sich oft gegenseitig –, dennoch sind sie voneinander trennbare Größen, die jeweils andere physische Entstehungsorte haben und auch getrennt voneinander beschrieben werden können<sup>16</sup>. Hierbei fließen sowohl auditive, als auch akustische und anatomische Erläuterungen in die Beschreibung mit ein.

#### 1) Temporale Merkmale:

Unter den temporalen Merkmalen versteht man im Allgemeinen die Sprechgeschwindigkeit im Sinne der „Realisierung bestimmter sprachlicher Einheiten (Laute/Silben/Wörter) in einer bestimmten Zeiteinheit [...]“ (Bose 2003: 43). Gemessen wird die Sprechgeschwindigkeit in Silben pro Sekunde, wobei sich als Vergleichsgröße aus phonetischen Messungen eine Durchschnittsgeschwindigkeit gesprochener Sprache von etwa fünf Silben pro Sekunde für deutsche Erwachsene ergab (vgl. z.B. Bose 2003: 141). Das auditiv wahrnehmbare Sprechtempo hängt auch von Pausenhäufigkeit, -länge und -dauer während des Sprechens ab. Oft korreliert der temporale Parameter mit dem artikulatorischen, indem bei erhöhter Sprechgeschwindigkeit die Artikulationsspannung abnimmt (vgl. Zilliken 1990: 17). Er ist im freien Sprechen auch abhängig von Sprechplanungsprozessen und der Sprechgliederung (vgl. Kowal 1991: 35 u. 70ff.).

#### 2) Dynamische Merkmale:

Unter Dynamik wird die wahrnehmbare Lautstärke und Lautstärkeänderung während des Sprechvorgangs verstanden. Sie bildet sich aus dem Anblasedruck und der

---

<sup>16</sup> Bose (2003: 39ff.) fasst hier die Ergebnisse verschiedener Autoren sehr übersichtlich zusammen, deshalb folge ich hier im Wesentlichen ihren Ausführungen (vgl. auch Geißner (1989: 76f.) und Zilliken (1991: 15ff.)).

Schwingungsweite der Stimmlippen und ist auf dem Oszillogramm in der Größe der Amplitude der aufgezeichneten Schwingung sichtbar. Das, was als Lautstärkevariation wahrgenommen wird, ist somit die momentane „Maximalauslenkung [...] von der potentiellen Ruhelage“ der Stimmlippen (Pétursson/Neppert 2002: 125). Daneben sind Lautheitswahrnehmungen unter anderem auch von der wahrnehmbaren Spannung abhängig (vgl. Bose 2003: 40). Eine erhöhte Lautstärke führt aufgrund des größeren Drucks oft zu mehr Spannung in den Stimmlippen und damit zu einer höheren Grundfrequenz, was sich in der Wahrnehmung des Parameters Melodik niederschlägt (vgl. Zilliken 1991: 18). Neben der mittleren Lautstärke werden hier der generell ausgeschöpfte Lautstärkebereich sowie die Häufigkeit der Lautstärkevariation mit in die Analyse einbezogen.

### 3) Melodische Merkmale:

Die melodischen Merkmale umfassen die Stimmtonhöhe sowie ihre Veränderung und Weite in Abhängigkeit zur Zeit. Die Tonhöhe richtet sich nach den Schwingungsperioden der Stimmlippen, welche von deren Spannung, anatomischer Beschaffenheit und dem subglottalen Druck abhängen (vgl. z.B. Bose 2003: 39).

Generell ergibt sich die Gesamtfrequenz der Stimme aus der Schwingungszahl der Stimmlippen pro Sekunde als Komplex aus mehreren Teilschwingungen, gemessen in Hertz (Hz). Als Grundfrequenz ( $F_0$ ) bezeichnet man dabei die messbare Periodenfrequenz der 1. harmonischen Teilschwingung. Sie bildet mit den oben erwähnten Kriterien die Basis für die auditiv wahrnehmbare Sprechstimmlage (vgl. Pétursson/Neppert 2002: 129, Kehrein 2002: 98f., Bose 2003: 40). Da die Tonhöhe somit ein auditives Komplexmerkmal darstellt, muss sie als solches auch analysiert werden. Hierbei wird die durchschnittliche Sprechtonhöhe relativ zur mittleren Sprechstimmlage (hoch – tief) betrachtet, in der die muskulären Spannungsverhältnisse ausgeglichen sind. Daneben fließen in die Melodik des Sprechausdrucks die Intervallgröße der Tonhöhenbewegung und die Frequenz des Tonhöhenbewegungswechsels mit ein<sup>17</sup>.

Die melodischen Merkmale spielen bei der Wahrnehmung des Stimmklangs eine sehr große Rolle, sowohl auf segmentaler als auch auf suprasegmentaler Ebene (vgl. Pétursson/Neppert 2002: 137). Denn Tonhöhenunterschiede werden bei Sprechern „sehr

---

<sup>17</sup> Geißner (1989: 76) fügt außerdem noch die Merkmale Klangfülle, Klangfarbe und Klangfarbenqualität hinzu. Diese werden unter dem Merkmal „Stimmklang“ hier jedoch gesondert behandelt.

genau wahrgenommen und sehr differenziert beurteilt (interpretiert) [...]“ (Zilliken 1991: 19).

#### 4) Artikulatorische Merkmale:

Die artikulatorischen Merkmale sind eng mit der muskulären Spannung des gesamten Mund- und Nasenraums und der Artikulationsorgane verbunden. Artikulatorisch wird ein Sprechausdruck zum einen unterschieden nach Art der generellen Lautrealisation. So wird eine sehr gespannte Artikulation als konsonantisch und eine eher ungespannte als vokalisch bezeichnet. Von der Artikulationsspannung hängt ebenfalls ab, ob die Artikulation als deutlich oder undeutlich wahrgenommen wird, und ob das Sprechen abgehackt wirkt oder Laute eher gebunden realisiert werden. Auch spielt die generelle Zungenlage im Ansatzrohr eine Rolle, wonach die Artikulation im gesamten eher vor- oder rückverlagert ist.

#### 5) Stimmklang:

Die generelle klangliche Eigenschaft der Stimme, d.h. die Art der Phonation und die Variation der Resonanz im Ansatzrohr, besitzt nach Auffassung verschiedener Autoren ebenfalls bedeutende Ausdrucksfunktion (vgl. etwa Fischer/Hartlieb 1984: 63, Eckert/Laver 1994: 64ff.) und kann die anderer Merkmale „vereindeutigen“ (Kranich 2002: 22). Auch hier handelt es sich um einen akustischen „Merkmalskomplex [...], in dem Frequenz, Amplitude und das Verhältnis der Partialtöne zum Grundton wirksam sind“ (Kranich 2002: 18). Subjektive Beschreibungen aus anderen Wirkungsbereichen sind nur unzureichend vergleichbar (vgl. Kranich 2002: 19f.). Eine erste anatomisch fundierte Klassifikation verschiedener ausdrucksrelevanter Stimmklangphänomene legte Trojan (1952) vor. Dazu fasste er klangliche Parameter wie „Überluft“ oder „Näselklang“ zu kleinsten akustischen Einheiten (sog. „Akuemen“) zusammen, wobei er Akuem als „Inbegriff aller Merkmale, durch die sich ein Gefühlszustand phonisch und artikulatorisch kund gibt“ definierte (1952: 13). Der hieraus entwickelte Merkmalskatalog war und ist für spätere Arbeiten wegweisend<sup>18</sup>.

Laver (1980) beschreibt das Zustandekommen dieser verschiedenen Variationen des Stimmklangs genauer anhand unterschiedlicher phonatorischer Settings. Bestimmend hierfür sind Modifikationen der Stimmlippenschwingung und die Weite der Glottisöffnung. Diese hängen wiederum eng mit den muskulären

---

<sup>18</sup> Vgl. etwa die Anmerkung in Fähmann (1960: 8f.) sowie Kranich (2002: 20).



Spannungsverhältnissen im Kehlkopf (Larynx) zusammen und werden deshalb auch als *laryngale Settings* bezeichnet. Drei voneinander unterschiedene Spannungen spielen hierbei eine Rolle: Die *adduktive Spannung* an den Arytenoidknorpeln, die Längsspannung der Stimmlippenmuskeln (*longitudinale Spannung*) und die Stärke der *medialen Kompression*, der muskulären Spannung am hinteren Teil der Glottis, die bestimmt, wie stark die Stimmlippen verschlossen sind<sup>19</sup>. Folgende Settings können nach Laver unterschieden werden:

- a) Die *Nullphonation* ist diejenige Einstellung, bei der kein Sprachschall hörbar wird, weil die Glottis weit geöffnet ist und die Luft aus der Lunge die Stimmlippen ungehindert passieren kann (Atemstellung). Dieses Setting wird zur Bildung stimmloser Laute genutzt (vgl. Kienast 2002: 20).
- b) Der *neutrale Phonationsmodus* (bei Laver: „Modal Voice“) bezeichnet die effektivste Art der Stimmgebung, die sogenannte Indifferenzlage der Grundfrequenz. Hier sind alle Spannungsverhältnisse ausgeglichen, die Stimmlippen schwingen periodisch und effizient, mit minimalem Druck- und Kraftaufwand, ohne dass Luft ungenutzt durch die Glottis entweicht und es zu hörbarer Geräuschbildung kommt. Diese Art der Stimmgebung stellt sozusagen die Norm dar, wovon alle anderen „Settings“ in einer bestimmten Art und Weise abweichen (vgl. Laver 1980: 94 u. 111).
- c) *Falsettstimme*: Hiermit wird eine Stimme mit extrem hoher Grundfrequenz im Vergleich zur Indifferenzlage des Sprechers bezeichnet. Diese kommt dadurch zustande, dass durch eine hohe longitudinale Spannung und extreme mediale Kompression nicht die gesamten Stimmlippen schwingen, sondern nur deren äußere Gewebeschichten hin zur Glottis. Dies führt zu schnellen Schwingungen mit sehr weit auseinander liegenden Obertönen, wodurch die Stimme nicht nur hoch, sondern auch sehr dünn wirkt (vgl. Laver 1980: 118ff.).
- d) *Behauchte Stimme*: Durch eine minimale adduktive Spannung und eine schwache mediale Kompression liegen bei dieser Art der Phonation die Stimmlippen nicht ganz so eng aneinander wie bei der neutralen. Die Glottis schließt deshalb nicht vollständig und lässt Luft entweichen, was als „Hauchen“ im Sprachschall hörbar wird. Auch schwingen die Stimmlippen weniger effizient als normal (vgl. Laver 1980: 134).

---

<sup>19</sup> Vgl. Laver (1980: 108f.) und die ausführliche Darstellung bei Kienast (2002: 19ff.), sowie Pétursson/Neppert (2002: 72ff.).

- e) *Flüsterstimme*: Durch eine hohe mediale Kompression ist der vordere Teil der Glottis geschlossen. Die adduktive Spannung ist jedoch sehr gering, die Aryknorpel sind nach außen gedreht und lassen in der hinteren Glottis eine kleine dreieckige Öffnung frei (das sogenannte „Flüsterdreieck“). Die hierdurch entweichende Luft wird als Turbulenzen hörbar und kann zur Bildung von Lauten genutzt werden, es entsteht jedoch keine Schwingung (vgl. Laver 1980: 121).
- f) *Knarrende Stimme*: Die Stimmlippen liegen eng aneinander und sind durch eine hohe mediale Kompression stark zusammengepresst. Die longitudinale Spannung ist jedoch sehr gering, wodurch die Stimmlippen dick und kompakt wirken. Gleichzeitig kommt aus der Lunge sehr wenig Druck, was dazu führt, dass die Stimmlippen unregelmäßig schwingen. Dies wird als knarrende Impulsanregungen hörbar (vgl. Laver 1980: 122f.).
- g) *Rauhe Stimme*: Insgesamt herrschen hier im Kehlkopf sehr hohe Spannungsverhältnisse an allen drei Punkten. Dies führt „zu einem unperiodischen Anregungssignal [...], das starken Mikroperturbationen hinsichtlich Amplitude und Frequenz unterworfen ist“ (Kienast 2002: 21). Diese unregelmäßige Grundfrequenz kann auch von untrainierten Hörern besonders gut wahrgenommen werden (vgl. Laver 1980: 127), wird jedoch häufig mit der knarrenden Anregung verwechselt (vgl. Scherer 1982: 197).

Diese Phonationsarten treten auch häufig in verschiedenen Kombinationen auf und können so komplexe Klänge ergeben (vgl. Laver 1980: 112). Zum Beispiel sind Rauheit und Behauchtheit zu dem Phänomen kombinierbar, das allgemein als Heiserkeit bekannt ist. Obwohl diese Arten des Stimmklangs auch pathologisch sein können, werden sie ebenso je nach Situation und Rolle des Sprechenden in der alltäglichen Kommunikation eingesetzt. Deshalb werden sie hier ebenfalls in die Analyse mit einbezogen (vgl. hierzu auch Bose 2003: 43).

Des Weiteren werden in dieser Arbeit unter dem Parameter Stimmklang diejenigen stimmlichen Eigenschaften gefasst, die sich auf den gesamten klanglichen Eindruck der Stimme auswirken. Genauer sind dies Klangfülle und Klangfarbe.

- h) *Klangfülle*: Als Klangfülle wird „die Zahl und Stärke der Harmonischen“ im Stimmklang bezeichnet (Fischer/Hartlieb 1984: 40, vgl. auch Bose 2003: 41). Je nachdem, wie die Artikulationsorgane zueinander positioniert werden, verändern sich die Resonanzeigenschaften des Ansatzrohres. So bewirkt zum Beispiel die

muskuläre Spannung des Rachens entweder einen vollen, obertonreichen, oder einen engen, obertonarmen Klang<sup>20</sup>. Aber auch die Stellung der Lippen und des Kehlkopfs sowie der Grad der Kieferöffnung spielen hierbei eine Rolle<sup>21</sup>.

- i) *Klangfarbe*: Dies bezeichnet nach Fischer und Hartlieb (1984: 41) den „Frequenzkomplex aus Grundton und harmonischen Obertönen nach Zahl und Stärke“. Je nachdem, welche Grundfrequenz zugrunde liegt und welche Obertöne verstärkt werden, kann eine Stimme entweder hell oder dunkel wirken. Das bedeutet, dass die Wahrnehmung der Klangfarbe auch oft „von der Tonhöhe beeinflusst“ ist (Bose 2003: 41). Dunkle Stimmklänge zeichnen sich zum Beispiel i.d.R. durch eine größere Energie im Bereich der niedrigeren Teilschwingungen aus (vgl. Bose 2003: 41). Dies hat nichts mit dem persönlichen „Timbre“ eines Sprechers zu tun (vgl. Fischer/Hartlieb 1984: 41), sondern ist ein Stimmcharakteristikum, das sich aus dem Zusammenspiel „von Zungenlage und Rachenweitung“ ergibt (Fährmann 1960: 60f.). Das „Timbre“ als „persönliche[r] Eigenklang“ (Fischer/Hartlieb 1984: 42) der Stimme entsteht wohl eher aus der „Energieverteilung im Stimmspektrum“ (Scherer 1982: 197).

#### 6) Sprechspannung:

Auffallend an den bisherigen Ausführungen ist, dass die muskuläre Spannung während des Sprechvorgangs in allen Parametern eine nicht unwesentliche Rolle spielt bzw. mit ihnen zusammenwirkt. Deshalb wird sie hier, in Anlehnung an Heilmann, als Merkmal gesondert betrachtet. Heilmann definiert Sprechspannung folgendermaßen:

„Die Autorin faßt „Sprechspannung“ als einen Oberbegriff auf, der sowohl Artikulationsspannung allgemein als auch Lautspannung im einzelnen umschließt, weiterführend die gesamte Muskulatur des Stimm- und Artikulationsraumes aber auch die zum Sprechen notwendige Körperspannung, verbunden mit Sprecherintentionalität“ (Heilmann 2002: 181).

Das bedeutet: Sprechspannung ist als Komplexmerkmal besonders wichtig für den Ausdruck der Stimme, weil Sprechen nicht ohne die oben beschriebenen muskulären Vorgänge ablaufen kann. Trotzdem werden von verschiedenen Autoren nur einzelne Spannungszustände wie die des oberen Ansatzrohres (vgl. Zilliken 1990: 15) oder der Rachenmuskulatur (vgl. Geißner 1989: 76) genannt, ohne hierfür ein solch nachvollziehbares einziges Merkmal einzuführen<sup>22</sup>. Dabei lässt sich anders die „ganzheitliche Komplexität“ (Heilmann 2002: 181), die der Hörer beim Sprecher durch

<sup>20</sup> vgl. Geißner (1989: 76). Trojan (1952: 104) hat dies als „faukale Enge bzw. Distanz“ bezeichnet und mit den Gefühlen für Ekel (viel Spannung) und Wohlbefinden (wenig Spannung) gleichgesetzt.

<sup>21</sup> Kienast hat die zugrundeliegenden Settings nach Laver sehr ausführlich beschrieben (2002: 21ff.).

<sup>22</sup> vgl. auch die Zitate verschiedener Autoren in Bose (2003: 46f.).

die Sprechspannung und ihre Kopplung an die anderen Parameter wahrnimmt, kaum darstellen. So bemerkt Krech (1987): „Die Sprechspannung [...] ist das grundlegende Mittel, mit dessen Hilfe der Ausdruck gestaltet wird. Auf ihr basiert der Gebrauch aller anderen [...] sprecherischen Mittel“ (zitiert nach Heilmann 2002: 182). Allein um „Sprechen als Gesamtprozess“ (Heilmann 2002: 182) darstellen zu können, kann man deshalb nicht darauf verzichten, das Merkmal Sprechspannung in eine auditive Analyse gesondert mit aufzunehmen.

### **2.3 Sprechausdruck und Persönlichkeitsbeurteilung – bisherige Untersuchungen**

Erste Versuche, bestimmte Eigenschaften eines Individuums und deren Wirkung aus seiner Stimme und Sprechweise heraus zu deuten, gab es schon in der Antike, um anschließend diese Sprechweisen in der Rede rhetorisch einsetzen zu können (vgl. Fährmann 1960: 3). Diese „vorwissenschaftliche“ und „intuitive“ Behandlung des Gegenstands hielt bis ins beginnende 20. Jahrhundert an. Auch aufgrund der oben erwähnten elementaren Erfassung des Sprechausdrucks durch Drach (vgl. Kapitel 2.2.2) gelang es danach erstmals, wissenschaftlich begründete Charaktertypen zu entwickeln und mit bestimmten Persönlichkeitseigenschaften in Beziehung zu setzen<sup>23</sup>.

Fährmann etwa bediente sich 1960 dazu den Stimmufnahmen einzelner Personen, sowie psychologischen Testverfahren zur Ermittlung von deren Persönlichkeitseigenschaften. Anschließend wurde das Sprachmaterial phonetisch, sprechphysiologisch und akustisch ausgewertet und zu den Ergebnissen der psychologischen Tests in Beziehung gesetzt. Leitend war hierbei die Frage, welche Sprechmerkmale auf welche Charakterzüge zutreffen. Anschließend ordnete er diese Stimmeigenschaften acht ermittelten Charaktersymptomgruppen zu. Dadurch fühlte er sich „doch in der Lage, einige für die praktische sprechdiagnostische Arbeit grundlegende Ergebnisse aufzuzeigen“ (Fährmann 1960: 154).

Für die weiteren Jahre stand vor allem die Frage im Vordergrund, wie andere Hörer in der Lage sind, die Persönlichkeitseigenschaften eines Sprechers aufgrund seines stimmlichen Ausdrucks zu beurteilen. Laut Addington (1968: 10) zeigte sich dabei untersuchungsübergreifend recht schnell, dass „the judges tended to stereotype the personality of the speaker according to the sound of his voice“. Dass es sich hierbei um stereotype, selten zutreffende Muster in der Beurteilung handelte, zeigten Vergleiche

---

<sup>23</sup> vgl. die vorgestellten Untersuchungen bei Fährmann (1960: 6ff.).

mit Selbsteinschätzungen der Sprecher. Dies führte Addington dazu, genau jenes stereotype Hörverhalten empirisch zu untersuchen, d.h. ob bestimmte stimmliche Merkmale tatsächlich hörerübergreifend bestimmte Einschätzungen bewirken (vgl. Addington 1968: 12). Dazu wurden simulierte Sprachaufnahmen verwendet, in welchen vier Studierende eines Stimmtrainingkurses dazu angehalten wurden, neun unterschiedliche „Stimmqualitäten“ bzw. Stimmklänge zu simulieren und diese in den Merkmalen Tempo und Tonhöhe zu variieren (vgl. ebd.: 13f.). Diese wurden anschließend in verschiedenen Zusammenstellungen von Hörergruppen anhand fester Adjektivskalen beurteilt (vgl. ebd.: 25). Als wichtigstes Ergebnis konnte festgehalten werden, dass einerseits Persönlichkeitsstereotype bestimmten stimmlichen Merkmalen mit großer Übereinstimmung zugeordnet werden können und zum andern – bezogen auf die untersuchten Parameter – „the vocal characteristics manipulated in this experiment were definitely effective in altering perceptions of personality“ (ebd.: 67; im Original kursiv).

Nichtsdestotrotz plädierte Scherer etwa ein Jahrzehnt später dafür, die Untersuchung des Zusammenhangs zwischen Stimme und Persönlichkeit trotz der „enttäuschende[n] Schlußfolgerung vieler früher Untersuchungen, [...] daß Persönlichkeitsurteile aufgrund von Stimmqualität auf zumeist irreführenden Stereotypen beruhten“ (Scherer 1982: 188), weiterzuführen. Dazu verwendete er in mehreren Studien erstmals „Proben interaktiver Sprache anstelle von Monologen“ (ebd.) und ließ nicht die Sprecher selbst ihre Persönlichkeit einschätzen, sondern legte Beurteilungen durch Bekannte und Freunde („peer-ratings“) zugrunde. Anschließend sollten amerikanische und deutsche Hörer die Stimmen kulturübergreifend persönlichkeitskonform einschätzen, was „mit überzufällig großer Wahrscheinlichkeit“ gelang (ebd.). Basis für die Untersuchungen waren Sprachproben simulierter Geschworenensitzungen, die zerschnitten und so angeordnet wurden, dass nur noch die Stimmqualität als Phänomen deutlich hörbar war (sog. „randomized splicing“, vgl. ebd.: 192). Insgesamt konnte Scherer zeigen, dass die Zuschreibung von bestimmten Persönlichkeitseigenschaften durch Sprache auch kulturübergreifend recht gut gelingt, wenn auch die Ergebnisse nur eingeschränkt valide sind (vgl. ebd.: 198ff.). Des weiteren kommen Scherer und Mitarbeiter zu dem Schluss, „dass weitere Forschungsarbeiten nötig [sind...], um diese ersten Hinweise auf die Wirkungen von Hinweisreizen auf die Personenwahrnehmung zu bestätigen“ (Scherer et al. 1982: 252).

Von Paeschke und Sendlmeier (1997) stammt eine Untersuchung, die in ihrem Gedankengang der hier vorgelegten am nächsten kommt. Die Autoren führten anhand verschiedener Hörereinschätzungen eine akustische und auditive Analyse der Redestile zweier Politiker eines Parteitags durch. Als Einschätzungsbogen diente dabei der Sprechausdruckskatalog von Krech. Zwar gelang es ihnen auf diese Weise, bei beiden Stimmen jeweils charakteristische Merkmale eines Redestils zu ermitteln, jedoch merken sie an, dass damit nicht abgesichert sei, „daß diese Merkmale auch die Ursache für den Mißerfolg von Scharpings Rede bzw. den Erfolg von Lafontaines Rede sind“ (Paeschke/Sendlmeier 1997: Abschn. 4).

Eckert und Laver (vgl. 1994: 152) weisen bei der Zusammenfassung verschiedener Untersuchungen zur Sprechwirkungsforschung darauf hin, dass das grundsätzliche Problem darin bestünde, die Stimme zum alleinigen Unterscheidungskriterium machen zu können und von Faktoren wie dem Sprachinhalt oder Alter und soziale Rolle des Sprechers zu trennen. Deshalb dürfe sich der Experte nie so unverblümt äußern, wie dies der Laie tue.

Genau aber diese „unverblühten Äußerungen“ sind nun erstmals Gegenstand einer vergleichenden Untersuchung, die daneben nicht nur isolierte Eindrücke, sondern einen gezielten Eindrucksvergleich beider Stimmen berücksichtigt. Dabei werde ich bei der Diskussion der Ergebnisse des Vergleichs auf relativ gesicherte Erkenntnisse der hier vorgestellten Untersuchungen zurückkommen.

## **2.4 Filmsynchronisation in Deutschland**

Schon seit den Anfängen des Filmgeschäfts bedient sich die Filmindustrie unterschiedlichster Methoden, um ihre Erzeugnisse auch einem fremdsprachigen Publikum zugänglich zu machen. Genügte hierfür im Stummfilmzeitalter das Übertragen der eingeblendeten Texttafeln in die jeweilige Landessprache, so machte spätestens die Einführung des Tonfilms komplexere Verfahren nötig (vgl. Maier 1997: 63ff.). Denn nun musste auch der über den auditiven Kanal vermittelte Inhalt übertragen werden, und zwar so, dass die parallele Rezeption von Bild und Ton dem Zuschauer keine oder nur wenig Schwierigkeiten bereitet und er der Filmhandlung folgen kann (vgl. Blaseio 2003: 163). Um dies zu erreichen, werden noch heute im Wesentlichen zwei Verfahren angewendet: die Bearbeitung eines Films durch Untertitel und Synchronisation. Synchronisation bezeichnet dabei im Allgemeinen „die Übersetzung

des ausländischen Originaldialoges in die deutsche Sprache sowie den technischen Vorgang der Verbindung des Originalbildstreifens mit dem [neuen] Tonstreifen [...]“ (Hesse-Quack 1969: 13). Dabei „leihen“ in diesem Fall deutsche Synchronsprecher den Schauspielern auf der Leinwand ihre Stimmen, indem sie lippensynchron den neuen Text zum Bild des Films einsprechen. Zwar gibt es auch Ansätze, die unter Synchronisation sämtliche Arten der nachträglichen Filmbearbeitung verstehen, wonach Untertitelung ebenfalls unter diesen Begriff fallen würde (vgl. Pruys 1997: 7), diese lassen jedoch außer Acht, dass Synchronisation und Untertitelung zwei völlig unterschiedliche Verfahrensweisen sind, sowohl in ihrer Bearbeitung, als auch in ihrer Wirkung (vgl. Blaseio 2003: 161f.). Denn der besondere Reiz der Synchronisation liegt ja gerade darin, dass hier vorhandene Körper mit „fremden“ Stimmen ausgestattet und mit ihnen als Ganzes wahrgenommen werden, also Original-Bild und neuer Ton „synchron“ laufen und gesehen bzw. gehört werden sollen. Deshalb wird auch in der vorliegenden Arbeit Hesse-Quacks Definition von Synchronisation als Abgrenzung zu anderen Verfahren beibehalten.

Während etwa in den skandinavischen Ländern ausländische Filme mit Untertiteln sehr erfolgreich liefen und laufen, zeigte sich in Deutschland schnell, dass diesem Verfahren hierzulande wenig Erfolg beschieden war (vgl. Maier 1997: 64). Ein breites Publikum lässt sich etwa seit den 1950er-Jahren nur durch synchronisierte Filme erreichen (vgl. Maier 1997: 68). Dass synchronisierte Filme hierzulande bis heute so populär sind, liegt laut Pruys (1997: 76) vor allem daran, dass „sie die Illusionswirkung einer Filmhandlung weitgehend aufrechterhalten können und am bequemsten aufzunehmen sind“<sup>24</sup>. Obwohl die Synchronisation von Filmen im Vergleich zur Untertitelung immer einen personellen, finanziellen und zeitlichen Mehraufwand bedeutet, rechnet sich diese Methode vor allem für amerikanische Filmverleiher besonders, weil Deutschland für die exportorientierte US-Filmindustrie einen der wichtigsten Märkte darstellt und dementsprechend bei großen Produktionen auch schon im Vorfeld finanziell mit einkalkuliert wird (vgl. Pruys 1997: 55). Auch bei den Filmen der Batman-Reihe stand deshalb eine Synchronisation außer Frage, zumal Filme wie „The Dark Knight“ als Verfilmungen bereits verlegter Comics nicht nur eine entsprechende Fangemeinde, sondern auch eine umfangreiche Marketing-Maschinerie hinter sich wissen. Daher trägt hierzulande die Synchronisation wesentlich zur Wirtschaftlichkeit eines Films bei.

---

<sup>24</sup> Neben einigen anderen Nachteilen des Untertitelungsverfahrens, vgl. hierzu die Ausführungen von Pruys (1997: 78f.), sowie Maier (1997: 41f.).

So ist es nicht verwunderlich, dass sich aus diesem Verfahren im Laufe der Jahre eine eigene Industrie mit ganz bestimmten Arbeitsmethoden entwickelt hat (vgl. Maier 1997: 71). Amerikanische Filmfirmen wie Warner gründeten dazu deutsche Verleihfirmen, die eng mit speziellen Synchronstudios in Deutschland zusammenarbeiten.

Diese erhalten neben einer Originalkopie des Films oft ein sogenanntes „International Tape“ (IT-Band), das nur die „Hintergrundatmosphäre“, also Geräusche und Musik, enthält, sowie ein Drehbuchexemplar mit Dialogen und Kameraeinstellungen. Auf Grundlage dieses Materials schreibt zunächst ein Dialogbuchautor ein möglichst lippen- und gestensynchrones Buch mit deutschen Dialogen, das die textliche Grundlage für die neue Tonspur bildet. Lippen- und gestensynchron bedeutet, dass der deutsche Text so formuliert wird, dass er später zur Mimik und Gestik der im Bild sichtbaren, sprechenden Schauspieler passt. Dazu gehören etwa sichtbare Artikulationsveränderungen sowie extraverbale Merkmale, die das Gesagte unterstreichen bzw. ergänzen. So müssen je nach Bildeinstellung alle Merkmale der Kommunikation zur Synchronität zwischen Ton und Bild beitragen. Bei Nahaufnahmen ist deshalb zum Beispiel stärker auf Lippensynchronität zu achten als in totalen Einstellungen.

Dieser neue Text wird dann in mehreren kleineren Einheiten (Takes) von einem Sprecher dementsprechend eingesprochen, sodass der Eindruck entsteht, die Stimme entspreche dem auf der Leinwand sichtbaren Körper. Denn oberstes Ziel der Filmsynchronisation ist die Erzeugung einer Einheit von Körper und Stimme zwischen Schauspieler und Synchronsprecher, die so nicht existiert. Hierbei kommt neben den obigen Kriterien der Stimme des Synchronsprechers ebenfalls eine entscheidende Rolle zu.

Es hat sich nämlich gezeigt, dass „das Stimmcasting bei noch nicht an StammsprecherInnen vergebenen SchauspielerInnen allein durch das Rollendesign der zu synchronisierenden Akteurinnen und Akteure bestimmt“ wird (Blaseio 2003: 163). Das heißt, der Sprecher wird nicht so ausgewählt, dass er und der synchronisierte Schauspieler möglichst ähnlich klingende Stimmen besitzen, sondern der Klang seiner Stimme soll beim Hörer die gleichen Eigenschaften transportieren, welche durch die äußere Erscheinung der Rolle hervorgerufen werden, um eben jene Illusionswirkung der Einheit von Bild und Ton noch besser zu erzeugen. Oder wie Bräutigam (2001: 29) es formuliert: Die Stimme „muss nicht mit dem zu sehenden Schauspieler übereinstimmen, sondern mit der Kunstfigur, die dieser verkörpert“. Dies erklärt auch die schon in



Kapitel 1 angedeuteten stimmlichen Unterschiede<sup>25</sup>. Ist nach diesen Kriterien einmal eine Stimme ausgewählt, so wird sie in der Regel für die weiteren Rollen des jeweiligen Schauspielers beibehalten, um die Illusion aufrechtzuerhalten, er würde mit eben genau dieser Stimme Deutsch sprechen (vgl. Maier 1997: 93). Denn auch auf Deutsch sollen diese Schauspieler unverwechselbar sein (vgl. Pruys 1997: 94; Maier 1997: 107). Ob und wie sich die beiden Stimmen unterscheiden, fällt dafür meist nicht ins Gewicht. Schwierig wird dieser Umstand nur, wenn der jeweilige Schauspieler plötzlich andere Rollen annimmt und seine deutsche, anhand der vorhergehenden Rollen ausgewählte Stimme sich nicht oder nur schwer ebenso verstellen lässt<sup>26</sup>. Dann kann es vorkommen, dass die Synchronstudios für die entsprechenden Schauspieler plötzlich einen anderen Sprecher mit „passenderen“ stimmlichen Eigenschaften verpflichten<sup>27</sup>. Oder aber, die Stimme wird beibehalten. Beides führt zur Irritation der Zuschauer (vgl. Maier 1997: 107f.), weil die Einheit zwischen („Rollen“)Körper und Stimme gestört bzw. weniger aufeinander abgestimmt scheint<sup>28</sup>.

Auch bei Heath Ledgers Rolle in „The Dark Knight“ war dies der Fall: Da Ledger in seinen ersten Filmen Jugendliche bis junge Erwachsene verkörpert hatte<sup>29</sup>, wurde wie üblich anhand des Rollendesigns als Stimme die eher hell und jugendlich klingende Stimme des Sprechers Simon Jäger ausgewählt, obwohl Ledgers eigene Stimme eher als dunkel und tief einzuordnen ist (vgl. die vorliegende Untersuchung, Kapitel 3.3.7). Die Rolle des psychopathischen Bösewichts Joker stellte nun ein für Ledger ungewöhnliches Rollenbild dar, dem sich die deutsche Stimme anzupassen hatte. Die Folge hiervon ist, dass die deutschen Zuschauer eine Diskrepanz zwischen Stimme und Rolle bzw. Original- und Synchronstimme stärker wahrnehmen, dies thematisieren und bewerten<sup>30</sup>. Ebenjene Bewertungen bilden die Basis für die vorliegende Untersuchung. Die Rolle des Jokers wäre somit ein Beispiel dafür, wie die Praxis der Synchronstudios,

<sup>25</sup> Vgl. auch die Anmerkungen Blaseios zu den Stimmen Thomas Magnums in der gleichnamigen Fernsehserie (vgl. Blaseio 2003: 163f.).

<sup>26</sup> Eine weitere Möglichkeit ist auch der Tod des Stammsprechers (vgl. Maier 1997: 107).

<sup>27</sup> So geschehen etwa bei dem US-amerikanischen Schauspieler George Clooney, der in seiner Rolle als Gangsterboss Danny Ocean in „Ocean's Eleven“ (USA 2001) und den Folgefilmen nicht mehr von Detlef Bierstedt, sondern von Martin Umbach synchronisiert wird, der mit einer wesentlich tieferen, dunkleren Stimme ausgestattet ist. Dies geschah auf Wunsch von Regisseur Steven Soderbergh, was auch zeigt, welche wichtige Stellung die Synchronisation ihrer Filme bei manchen ausländischen Regisseuren und Produzenten hat (vgl. <http://www.jungstars.net/martin-umbach/vita.htm> (05.08.09), sowie Bräutigam 2001: 27).

<sup>28</sup> Laut Bräutigam (2001: 28) war dies nicht immer so, ein mögliches Zeichen dafür, dass die deutschen Stimmen hinter den Schauspielern in den letzten Jahren mehr in den Fokus der öffentlichen Wahrnehmung rücken (vgl. Blaseio 2003: 161).

<sup>29</sup> „Seine ersten Rollen waren die von Sunnyboys, Helden und Rittern [...]“ (Pitzke 2008). Vgl. auch die Filmographie in der „Internet Movie Database“ <<http://www.imdb.de/name/nm0005132/>> (05.07.09).

<sup>30</sup> Vgl. etwa den Kommentar von LuckyLenny92 auf der Liste der Userkommentare im Anhang.

einen Sprecher nach der jeweiligen Rolle und nicht nach stimmlicher Ähnlichkeit zum Schauspieler auszuwählen, die Wahrnehmung und Wirkung einer Figur für den deutschen Zuschauer verändert. Ob dies wirklich zutrifft, soll im Folgenden gezeigt werden.

### 3. Ein Vergleich zwischen Originalstimme und Synchronisation

#### 3.1 Sprechausdruck im synchronisierten Film – Vorüberlegung, Hypothesen

Fasst man die bisherigen Ausführungen zusammen, so ist festzuhalten, dass ein Unterschied der Stimmen zwischen Originalfassung und synchronisiertem Film durchaus ebenso andere stimmliche Eindrücke und Wirkungen beim Zuschauer bzw. Hörer hervorrufen kann, weil dieser stimmliche Eindruck auch in Schauspiel und Synchronisation durch bestimmte Merkmalsausprägungen entsteht, die im Sprechausdruck vermittelt werden. Denn die verschiedenen Eigenschaften einer Rolle müssen vom Schauspieler neben der Manifestation im körperlichen „Verhalten“ auch stimmlich hörbar gemacht werden; diese versucht der Synchronsprecher dann mit seiner Stimme erneut zu transportieren, wenn die Rolle ähnlich angelegt werden soll<sup>31</sup>. Dabei ist davon auszugehen, dass diese stimmlich ausdrucksimmanenten Eigenschaften nicht durch bewusste sprechorganische Modulationen des Schauspielers bzw. Sprechers im Sprechschall hörbar werden. Dies wäre von beiden nicht zu leisten und eine „bewusst-reflektierte, individuelle Gestaltung“ (Stock/Suttner 1991: 72) würde außerdem gekünstelt wirken. Schauspieler und Synchronsprecher – welche ja meist ebenfalls gelernte Akteure sind (vgl. Bräutigam 2001: 25ff.) – agieren so, als würden sie in diesem Moment genau so spontan denken, fühlen und sprechen. Damit erzielen sie eine bestimmte Wirkung beim Publikum, das diese bestimmten Zuständen zuordnen kann (vgl. Kranich 2002: 75). Kranich (2002: 76) geht mit dem Hinweis auf den Theaterpädagogen Stanislavsky davon aus, dass ein Schauspieler die darzustellende Rolle dazu in ihrem momentanen Zustand genauso von innen heraus durchleben müsse, wie ein in der gleichen Situation spontan agierender Mensch. Er nimmt dabei „den gesamten Körper einschließlich Stimm- und Sprechapparat ins ‚Schlepptau‘, wodurch die charakteristischen phonetischen Muster [...] ‚wie von selbst‘ entstehen“ (vgl. auch Stock/Suttner 1991: 69). Verschiedentlich wurde schon ausgeführt, dass Zuschauer/-hörer eine vorhandene bzw. fehlende Rollenkonformität der Stimme durch diese Merkmale durchaus wahrzunehmen in der Lage sind, wenn sie sich in der genauen Beschreibung auch schwer tun (vgl. Kapitel 2.4, Eckert/Laver 1994: 44).

---

<sup>31</sup> Zu einem gegenteiligen Fall vergleiche die Ausführungen anhand von „Rambo“ in Blaseio (2003: 166ff.).

Daraus folgt, dass sich vorhandene Eindrucksunterschiede auch in den jeweiligen Sprechausdruckparametern zeigen lassen müssten. Dem versucht die vorliegende Analyse nachzugehen. Folgende Hypothese wird hierzu formuliert:

H<sub>1</sub>: Durch eine akustische und auditive Analyse der Stimmen Heath Ledgers und des Synchronsprechers Simon Jäger im Film „The Dark Knight“ ist es möglich, deren Parameter auf den subjektiven Eindruck von Hörerurteilen zu beziehen und diesen so nachvollziehbar zu machen.

Voraussetzung für das Zutreffen der Hypothese ist allerdings, dass zwischen beiden Stimmen auch deutliche Unterschiede bestehen.

Die Betrachtung einzelner Parameter ist hier in sofern von Interesse, als dass sich womöglich zeigen ließe, wie sehr ein Höreindruck von der Veränderung einzelner Parameter abhängt bzw. wie differenziert Hörer minimale stimmliche Unterschiede mit verschiedenen Eindrücken verbinden.

Bewertungs- und Datengrundlage, sowie Vorgehen und Ergebnisse dieser Analyse sollen nun im folgenden erläutert werden.

## **3.2 Material und Methoden**

### **3.2.1 Die Vergleichgrundlage: Heath Ledger als Joker in „The Dark Knight“**

Sprechsprachliche Äußerungen, wie die Heath Ledgers in der Rolle des Jokers im Film „The Dark Knight“, geschehen immer in einem bestimmten historischen und situativen Kontext. Ebenso erfolgt ihre Perzeption immer durch in einer bestimmten Weise sozialisierte Individuen in einer Hörsituation, von Drach bereits definiert als „Gesamtheit aller vorausliegenden psychischen Erlebnisse bis zum Augenblick des Hörens [...]“ (Drach 1926, zitiert nach Geißner 1988: 77). Diese Erfahrungen des Hörers wirken sich auch auf die Interpretation der verschiedenen Elemente dieser sprachlichen Äußerung aus<sup>32</sup>. „Reine“ Sprachrezeption nur auf Grund des dargebotenen Schallereignisses gibt es nicht. Deshalb ist die Entstehung von Sprechwirkung immer situationsabhängig und auch als solche zu analysieren und zu verstehen (vgl. Stock 1991: 46). Die subjektiven Erfahrungswerte und das genaue Wissen der Hörer bleiben dabei allerdings dem Auge des Beobachters verborgen. Er kann nur versuchen, diese durch den Nachvollzug der äußeren Hörsituation beschreibbar zu machen (vgl. Stock

---

<sup>32</sup> Vgl. auch die Ausführungen in Kapitel 2.2.1, sowie Stock (1991: 27f.).

1991: 47). Hierzu zählen der zeitliche Rahmen, in dem die Äußerung dargeboten wird, aber auch allgemein bekanntes Wissen über Hintergründe des Sprechers, die zugrundeliegende Rollen- und Figurenkonstellation und die mögliche Sicht des Hörers auf diesen Situationskomplex aufgrund bekannter Fakten. Deshalb sollen im folgenden die entsprechenden Merkmale des in dieser Arbeit verwendeten Sprachmaterials kurz skizziert werden.

Das Figureninventar, dessen Geschichte im Film „The Dark Knight“ erzählt wird, wurde ab den 1930er-Jahren von einer Gruppe aus Zeichnern und Autoren um den amerikanischen Comicautor Bob Kane erdacht<sup>33</sup>. Titelgebend ist dabei die Figur des Batman, eines selbsternannten Rächers im Fledermauskostüm, der es sich zur Aufgabe gemacht hat, die Verbrecher seiner Heimatstadt Gotham City zu bekämpfen. Hinter Batman steht Bruce Wayne, ein junger Milliardär, der als Kind die Ermordung seiner Eltern mit ansehen musste und durch die Erschaffung seines Alter Egos einen Weg gefunden hat, dieses traumatische Erlebnis zu verarbeiten. Neu an dieser Figur war zur damaligen Zeit, dass Batman als Superheld ganz ohne Superkräfte auskommt und sich allein auf seine Intelligenz, detektivischen Spürsinn, körperliche Kraft und Ausdauer sowie zahlreiche technische Spielereien verlässt, was den Geschichten einen „realistischeren“ Anstrich verlieh. Daneben erschufen die Autoren im Laufe der Zeit zahlreiche Gegner, die Wayne auf diese Weise zu bekämpfen hatte. Eine davon ist die Figur des Jokers, eines gescheiterten Unterhalters, der durch einen Unfall in einer Chemiefabrik mit einem irreversiblen Clownsgesicht und einem diabolischen Grinsen ausgestattet ist<sup>34</sup>. Daraufhin wird er zum zynisch-sarkastischen Psychopathen, der durch makaberen Schabernack in Gotham City Angst und Schrecken verbreitet. Sein Markenzeichen ist die Joker-Spielkarte, was seinen spielerisch-albernen Charakter unterstreicht (vgl. auch Dath 2005: 8).

Da dieser Comicstoff bereits mehrmals verfilmt wurde, war auch die Figur des Jokers schon in unterschiedlicher Weise auf Leinwand und Bildschirm zu sehen. Dabei ist hier besonders anzumerken, dass die Interpretation der Rolle je nach Film ganz unterschiedlich ausfiel. Bis zum Film „The Dark Knight“ war Kinozuschauern vor allem die Darstellung des Schauspielers Jack Nicholson im Gedächtnis, der die Rolle des Jokers im Film „Batman“ des Regisseurs Tim Burton ganz im Sinne der Comicvorlage auslegte und damit dem Film zum Erfolg verhalf. Christopher Nolan legte nun vor diesem Hintergrund in seinem Film eine recht neue Interpretation der

---

<sup>33</sup> vgl. hierzu auch bezüglich der genauen Figurenentwicklungen Dath (2005) und Banhold (2009).

<sup>34</sup> vgl. hierzu den Comic „The Killing Joke“ in Klage et al. (Hgg. 2005), S. 111-156.

Figur vor: Die Elemente des Comics und der Überzeichnung fehlen, ansonsten wird aus dem Joker ein nachlässig geschminkter, namenloser Terrorist, der in seinen Handlungen meist nichts Schalkhaftes mehr an sich hat. Sein einziges Ziel ist es, die Welt ins Chaos zu stürzen. Diese Rollenveränderung wurde auch immer wieder in der öffentlichen Rezeption des Films thematisiert. So schreibt etwa die Redakteurin Christiane Peitz (2008) in ihrem Filmkommentar im „Tagesspiegel“:

„War Jack Nicholsons Joker im 1989er-„Batman“ ein Künstler des Todes, so ist Heath Ledger der Gott des Chaos. Bisher waren Batmans Gegner Artisten: Nicholson als Image-Designer und Actionpainter, dessen Visage ein Produkt toxischer Chemie war und der sich mit vergifteten Kosmetika für die e [sic!] Demütigung rächte. [...] Sein [Heath Ledgers, C.H.] Joker hat mit Ästhetik so wenig zu tun wie mit Rache“.

Die Transportierung dieses neuen Rollenbildes schlägt sich auch in der Wahrnehmung des stimmlichen Ausdrucks nieder, wie etwa die Kommentare auf YouTube zeigen<sup>35</sup>.

Eine weitere Besonderheit, die der Rolle des Jokers in „The Dark Knight“ zusätzlich Aufmerksamkeit verlieh, war die Tatsache, dass Heath Ledger mit ihr sein traditionelles Rollenschema des Sunnyboys ins Gegenteil verkehrte. Dementsprechend deutlich wurde auch der stimmliche Unterschied zu diesen Rollenbildern wahrgenommen<sup>36</sup>.

All diese situativen und historischen Tatsachen spiegeln sich laut der Kommentare auch in der Wahrnehmung der Figur durch die Hörer wider, welche die stimmliche Leistung Heath Ledgers und Simon Jägers danach interpretieren und beurteilen. So wird einerseits angemerkt, dass die Originalstimme im Gegensatz zur deutschen Stimme nicht zum bisherigen Rollenbild passe, zum anderen wird gerade der Synchronstimme wenig Kongruenz zum jetzigen Rollenbild attestiert. Hierauf wird bei der Analyse und dem Vergleich beider Stimmen noch genauer einzugehen sein.

Abschließend sei hier noch kurz die essentielle Rahmenhandlung des Films „The Dark Knight“ skizziert, um die einzelnen Szenen der Analyse anschließend besser in den situativen Filmkontext einbetten zu können (vgl. Banhold 2009: 88ff.).

Nachdem der Vorgängerfilm „Batman Begins“ erzählte, wie Bruce Wayne mit der Intention, das Verbrechen zu bekämpfen, sein Alter Ego Batman erschuf und damit eine Interpretation der Figur jenseits des Comichaften etablierte, hat Batman in „The Dark Knight“ inzwischen fast alle Verbrecher dingfest gemacht. Hilfe bekommt er hierbei vom Polizisten Lt. Jim Gordon und dem neuen Staatsanwalt Harvey Dent, der als „Weißer Ritter“ gefeiert wird und das Verbrechen endlich jenseits von Batmans

<sup>35</sup> vgl. etwa die Kommentare der Nutzer TimoRac, Friesa360 und vigeous auf der Liste der Userkommentare im Anhang.

<sup>36</sup> Vgl. die Ausführungen in Kapitel 2.4.

Untergrundmethoden auf legalem Wege bekämpfen soll. Doch hier taucht plötzlich der Joker auf, ein namenloser Terrorist mit Clownsgesicht, der den Mafiabossen der Stadt anbietet, Batman zu töten, wenn sie ihm die Hälfte ihres Vermögens überlassen. Durch Batmans Aktivitäten arg in Bedrängnis gebracht, willigen diese ein und der Joker befiehlt per Videobotschaft, Batman solle seine wahre Identität preisgeben, ansonsten werde jeden Tag ein Mensch sterben. Doch Batman weigert sich aufzugeben, kann so aber den Tod von fünf Menschen durch den Joker nicht verhindern. Als die Öffentlichkeit ein Versteckspiel Batmans deshalb nicht mehr hinnehmen will, behauptet der Staatsanwalt Harvey Dent, Batman zu sein, und kann so den Joker in eine Falle locken und ihn festnehmen, ohne dass Bruce Wayne als Batman aufgeben muss. Doch der Joker hat inzwischen die ganze Stadt mit seinen Leuten infiltriert und organisiert aus dem Gefängnis heraus die Entführung Harvey Dents und seiner Freundin Rachel Dawes, um Batman weiter unter Druck zu setzen. Im Verhör mit Batman macht er dann außerdem klar, dass sein wahres Ziel darin besteht, durch die Verbreitung von Panik und Chaos dafür zu sorgen, dass die Bürger Gotham Citys ihre Moral verlieren. Indem er anschließend Rachel Dawes töten lässt, gelingt es ihm, Harvey Dent so zu demoralisieren, dass dieser nun an den vermeintlichen Drahtziehern der Verbrechen Rache übt. Der Joker kann währenddessen fliehen und droht in einem letzten Coup, zwei Passagierschiffe in die Luft zu sprengen. Batman kann Harvey Dent und den Joker allerdings im letzten Moment stellen und so weitere Morde verhindern.

### 3.2.2 Auswahl, Gewinnung und Verteilung der Hörerurteile

Wie bereits in Kapitel 1 und Kapitel 2.4 angedeutet, dienen in der hier vorgenommenen Untersuchung als Grundlage der vergleichenden Wirkungsbeschreibung des Sprechausdrucks Eindrücke, die Nutzer im Internet abgegeben hatten. Dabei handelt es sich um Kommentare zum Video „The Dark Knight Synchro-Vergleich (Englisch/Deutsch)“ auf dem frei zugänglichen Videoportal YouTube.com<sup>37</sup>.

Dies ist eine Internetseite, auf welcher registrierte Nutzer kostenlos Videos online stellen können, die dann jedem Internetnutzer zur Betrachtung zugänglich sind<sup>38</sup>. Andere registrierte Nutzer können unter ihrem User- oder Nickname (vgl. Runkehl/Schlobinsky/Siever 1998: 74) ein Video bewerten und mit einem Kommentar

---

<sup>37</sup> URL: <http://www.youtube.com/watch?gl=DE&hl=de&v=kW3KP3NgIBs> (21.07.09).

<sup>38</sup> Abgesehen von Videos, die als nicht jugendfrei klassifiziert sind. Diese können nur von registrierten Nutzern gesehen werden.

versehen. Diese Kommentare werden unterhalb des Fensters, in welchem das Video zu sehen ist, chronologisch aufgelistet, wobei der jüngste Kommentar jeweils ganz oben platziert wird. Die Kommentare äußern sich häufig zum Inhalt des Videos oder geben die Bewertung durch den Nutzer in ausführlichen Worten wieder. Gelegentlich erhält diese Funktion auch den Charakter eines Internetforums oder einer Mailingliste, denn es kann auch auf andere Einträge direkt geantwortet und Bezug genommen werden (vgl. hierzu Franz 2008: 18ff.). So können ganze Kommentarserien zu einem bestimmten Thema entstehen, die jedoch nicht, im Gegensatz zu Forum oder Mailingliste, thematisch geordnet sind, sondern in Reihe chronologisch untereinander stehen.

Das Video „The Dark Knight Synchro-Vergleich (Englisch/Deutsch)“ hat nun den „Vergleich zwischen der Original-Version einer Szene aus 'The Dark Knight' und der deutschen Synchronisation“ zum Thema, wie der verantwortliche Nutzer „GermanDarkKnight“ in der Beschreibung des Videos mitteilt. Der Vergleich erfolgt mittels eines Zusammenschnitts der Szene ab TC 47:13<sup>39</sup>, indem einzelne Äußerungen des Jokers auf Englisch und Deutsch jeweils hintereinandergeschnitten werden. Bemerkenswert ist, dass das Video bereits am 23. Juli 2008 hochgeladen wurde, also knapp einen Monat vor dem offiziellen Kinostart am 21. August 2008. Trotzdem sind die dargestellten Szenen auch in ihrer Synchronisation mit denen des fertigen Films identisch. Bisher<sup>40</sup> gaben über 300 registrierte Nutzer einen Kommentar dazu ab, in welchem sie sich meist zum für sie hörbaren Unterschied der dargestellten Stimmen und Sprechweisen äußerten. Dabei fällt auf, dass 46 der 320 Kommentare nicht nur eine allgemein qualitative Bewertung über die Leistung beider Sprecher enthalten, sondern gleichzeitig einen unmittelbaren Eindruck der Stimmen auf den jeweiligen Nutzer wiedergeben. Dies führte zu der Frage, ob diese Eindrücke völlig willkürlich abgegeben werden, oder ob sich doch Gemeinsamkeiten in Bezug auf beide Stimmen ergeben. Genauer: Gibt es tatsächliche Übereinstimmungen dieser Eindrücke, die Nutzer in Bezug auf eine oder beide Stimmen im Internet schildern, obwohl sie von völlig unterschiedlichen Individuen zu verschiedenen Zeiten unabhängig abgegeben wurden? Nur so erscheint es sinnvoll, diese konformen Eindrücke in einer stimmlichen Analyse nachvollziehbar machen zu wollen.

---

<sup>39</sup> Der *Timecode* (TC) gibt jeweils den Start- bzw. Endzeitpunkt einer Szene der DVD „The Dark Knight“ in Stunden, Minuten und Sekunden an, auf welche im Text referiert wird. Er zeigt die beim Abspielen der DVD sichtbare Zeitählung. So sind die entsprechenden Szenen problemlos im Suchlauf aufzufinden (vgl. auch Pruys 1997). Zum Abspielen der DVD am PC wurde hier das Programm WinDVD® verwendet.

<sup>40</sup> Stand: 21.07.09.



Selbstverständlich wurden auch in anderen Foren die Auswirkungen der Synchronisation auf die Wirkung der Figur diskutiert<sup>41</sup>. Dennoch wurden aus den folgenden vier Gründen die Kommentare dieses YouTube-Videos als Wirkungsgrundlage gewählt:

1. Alle Kommentare beziehen sich inhaltlich auf den hörbaren Unterschied zwischen der Stimme Heath Ledgers und seiner Synchronisation – ob der Intention, mit der das Video ins Netz gestellt wurde. Sie sind somit thematisch in sich einheitlich, was in allgemeineren Foren nicht immer der Fall ist.
2. Durch die Tatsache, dass die Kommentare direkt an eine Vergleichsgrundlage in Form des Videos angeschlossen sind, ist davon auszugehen, dass die geschilderten Eindrücke mit hoher Wahrscheinlichkeit nicht nur auf „Hörensagen“ beruhen, sondern vielmehr aufgrund der vorherigen Betrachtung zumindest dieses Videos abgegeben

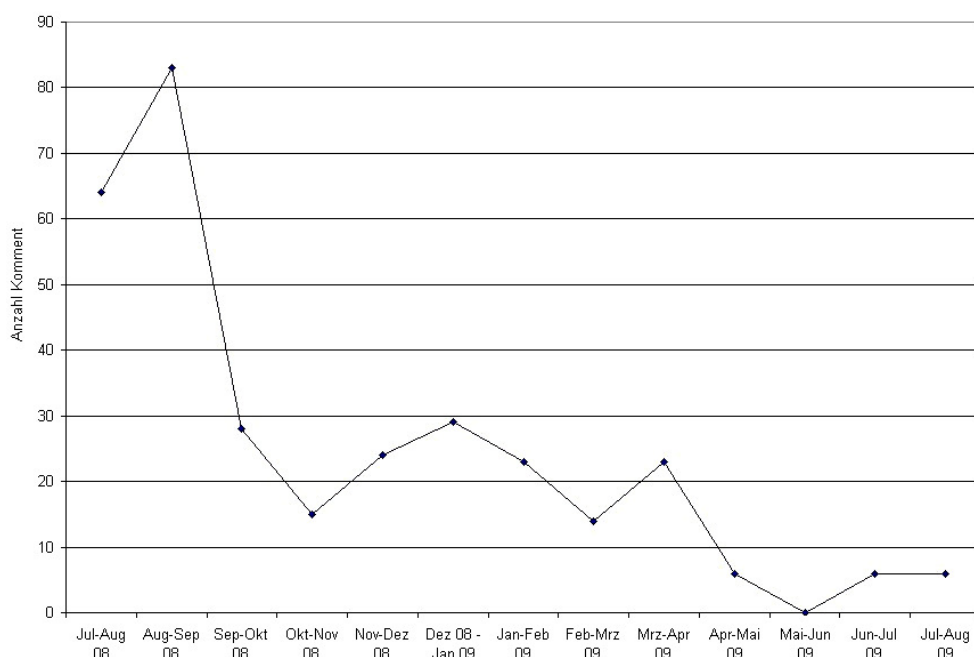


Abb. 3.1: Anzahl der Kommentare in Relation zum Abgabezeitraum.

wurden und somit der Eindruck recht unmittelbar sein dürfte. Diese Schlussfolgerungen wäre bei Äußerungen in anderen Foren nicht ohne weiteres zulässig, da hier der unmittelbare zeitliche Bezug zwischen Eindruck und Kommentar nicht unbedingt gegeben und nur schwer nachzuvollziehen ist. Zumal sich manche Eindrücke hier auf den Trailer beziehen können, dessen Synchronisation nur eine vorläufige Version des Endprodukts ist. Untermuert wird diese Annahme, wenn man sich die Abgabe der Kommentare in

<sup>41</sup> Vgl. etwa die Kommentare der User „Matze“ und „Apcher“ auf <http://www.trailerseite.de/archiv/trailer-2008/batman-the-dark-knight-trailer.html> (05.08.09)

Relation zur Zeit betrachtet (vgl. Abbildung 3.2 sowie im Anhang S. X). Bezogen auf den gesamten Zeitraum vom Hochladen des Videos bis zum Abschluss der Untersuchung am 21. Juli 2009 fällt auf, dass es drei Hauptgipfel gibt, an welchen am meisten Kommentare zu besagtem Video abgegeben wurden. Das globale Maximum fällt in den Zeitraum August/September 2008, also in unmittelbarer zeitlicher Umgebung zum deutschen Kinostart am 21. August 2008. Hier wurden, höchstwahrscheinlich nach Betrachtung des Films, überdurchschnittlich viele Kommentare abgegeben. Ein weiteres lokales Maximum wird im Zeitraum Dezember/Januar sichtbar, in welchen das Erscheinen der DVD des Films am 21. Dezember 2008 fällt. Der nächste Gipfel im März 2009 ist womöglich durch die geschenkrächtigen Osterfeiertage zu erklären. Das heißt, durch diese zeitliche Verteilung ist es ebenfalls mehr als wahrscheinlich, dass die Nutzer vor allem relativ zeitnah zur Betrachtung des Films entsprechende Kommentare abgaben<sup>42</sup>.

3. Auch wenn sich einige Kommentare meist nur auf eine der beiden Stimmen beziehen, ist mit insgesamt 46 verwendbaren Kommentaren zur Stimmwirkung eine zumindest ausreichende Datengrundlage an vergleichenden Eindrücken gegeben. Zumal die Eindrücke häufig von sich aus im Komparativ abgegeben wurden. Somit schien hier ein thematisch gut eingrenzbare, ausreichend großes Urteilskorpus an spontan entstandenen Eindrücken gefunden.

4. Die Nutzer, welche ihr Alter angaben, sind als Gruppe bezüglich dieses Merkmals recht homogen. Nur 5 von 34 sind älter als 25 und nur vier von 34 jünger als 18 Jahre bei einem Durchschnittsalter von ungefähr 22 Jahren. Allerdings ist die Einschränkung zu machen, dass sich diese Angaben nur schwer auf ihre Richtigkeit hin überprüfen lassen.

Um die Konstanz bestimmter Aussagen zu prüfen bzw. deren Willkür auszuschließen, wurden die individuell beschriebenen stimmlichen Eindrücke nach gemeinsamem semantischem Gehalt in Eindrucksgruppen zusammengefasst. Folgende Gruppen und Häufigkeitsverteilungen ergaben sich dabei<sup>43</sup>:

---

<sup>42</sup> Dies wird durch explizite Formulierungen in den Kommentaren wie: „Hab gestern den Film gesehen.“ noch bestätigt.

<sup>43</sup> Vgl. die wortwörtlichen Kommentare im Anhang.

<b>Eindrucksgruppen</b>	<b>Original</b>	<b>Synchronisation</b>	<b>Gesamt</b>
verrückt/wahnsinnig/irre/Psycho	2	23	
komisch/albern/witzig	0	3	
böse/fies/brutal/drastisch	4	2	
furchteinflößend/gefährlich/gruselig	5	1	
tief	3	0	
rauh	3	0	
<b>Gesamt</b>	<b>17</b>	<b>29</b>	<b>46</b>

Tabelle 3.1: Häufigkeitsverteilung der Hörereindrücke

Deutlich wird hier bereits, dass, obwohl insgesamt mehr Urteile für die Synchronstimme abgegeben wurden, ein deutlicher Unterschied in der Beurteilung der beiden Stimmen zu erkennen ist. So schreibt allein die Hälfte aller abgegebenen Urteile der Synchronstimme die Merkmale der Gruppe *verrückt* eher zu, während dieser Eindruck bei der Originalstimme nur in zwei von 46 Kommentaren überwiegt. Auch wird der Eindruck *komisch/albern/witzig* lediglich der Synchronstimme zugeordnet. Daneben existieren mit *tief* und *rauh* zwei Stimmklangsbeschreibungen, die laut Urteil nur auf die Originalstimme zutreffen. Es scheint also, dass die beiden Stimmen bei allen 46 Hörern einen unterschiedlichen Eindruck hinterlassen haben, denn kein Merkmal ist gleich dominant, im Gegenteil, bei *verrückt* gibt es sogar einen Wahrnehmungsunterschied von 23:2. Das heißt nicht, dass sich beide Stimmen völlig unterscheiden müssen<sup>44</sup>, denn die Eindrücke bestehen, wie oben bereits angedeutet, größtenteils aus komparativen Bewertungen, also einem „eher“ oder „mehr“. Nichtsdestotrotz geht hieraus hervor, dass bestimmte stimmliche Eigenschaften beiden Stimmen unterschiedlich stark zugeordnet werden, die Stimmen sich in ihrem vergleichenden Eindruck somit anscheinend recht klar voneinander unterscheiden.

Hervorzuheben ist, dass es sich hierbei um recht spontane Eindrücke handelt, die nicht in einem Beurteilungstest unter Laborbedingungen abgegeben wurden. Der Vorteil dieser Methode ist, dass störende und die Beurteilung der Sprechwirkung verzerrende Faktoren wie Beobachterparadoxon und Versuchsleiterereffekt entfallen, weil die Beurteilenden zum Zeitpunkt ihrer Bewertung gar nicht wissen konnten, dass ihre Einschätzungen Grundlage einer Analyse sein würden (vgl. Stock 1991: 53f.). Die Nachteile wie unbekannte Hörsituation oder nicht konforme Darbietung des

<sup>44</sup> Der Eindruck „sarkastisch“ wurde von einem User z.B. beiden Stimmen zugeordnet, jedoch ist hier nicht ganz ersichtlich, ob er sich auf die Sprechweise oder den Inhalt des Gesagten bezieht, deshalb fehlt dieser Eindruck in der Auflistung.

Sprachmaterials sind dennoch leider ebenso offenkundig und werden in der Diskussion der Ergebnisse (Kapitel 4) aufzugreifen sein.

Außerdem lassen sich die Hörerurteile schwer einer statistischen Beschreibung zugänglich machen, da es sich um reine Häufigkeitsangaben handelt. Es kann somit immer nur von absoluten Häufigkeitsanteilen eines Eindrucks beider Stimmen in Bezug auf die Gesamtzahl der Kommentare ausgegangen werden. Hierdurch lassen sich allerdings wohl immerhin gewisse Tendenzen der Sprechwirkung aufzeigen. Inwieweit diese Tendenzen jedoch tatsächlich eine stimmliche Grundlage haben, soll nun analytisch direkt am Sprachmaterial des Films überprüft werden.

### 3.2.3 Auswahl und Gewinnung des Sprachmaterials

Als Vergleichsgrundlage von Original- und Synchronstimme des Jokers dienten sprachliche Äußerungen der Figur im Film „The Dark Knight“, sowohl in der englischen Originalfassung als auch auf der deutschen Synchrontonspur. Leider war es aus technischen und rechtlichen Gründen nicht möglich, diese beiden Tonspuren isoliert zu bekommen. Deshalb wurde ein Weg gefunden, das sprachliche Material von der DVD des Films zu gewinnen. Vorteil dieser Methode ist, dass auf einer DVD sowohl die deutsche als auch die englische Tonspur in digitaler Qualität vorliegen und separat abgespielt werden können. Der Nachteil besteht darin, dass nur der komplette Film zur Verfügung steht, das heißt, auf beiden Tonspuren sind neben den Dialogen sowohl Geräusche als auch Filmmusik enthalten. Um jedoch – wie für die vorliegende Untersuchung geplant – eine akustische Analyse einzelner Parameter durchführen zu können, durfte das Sprachmaterial zur Berechnung im Computer nicht durch solche Nebeneffekte überlagert sein. Dies führte dazu, dass sich das Korpus der möglichen sprachlichen Äußerungen drastisch reduzierte<sup>45</sup>. Denn nahezu alle längeren Äußerungen des Jokers sind im Film mit mehr oder weniger laut wahrnehmbarer Musik unterlegt oder durch Geräusche unterbrochen. Hinzu kommt noch die Tatsache, dass dieser „Hintergrundsound“ auf beiden Tonspuren unterschiedlich abgemischt und daher verschieden laut zu hören ist, das heißt, die Stimmen kommen somit auch

---

<sup>45</sup> vgl. zu diesem Problem auch Braun/Heilmann (2006: 170). Bei ihrer Analyse von Emotionen im synchronisierten Film umgingen die Autorinnen diese Schwierigkeit, indem sie als Materialgrundlage eine synchronisierte, amerikanische Fernsehserie verwendeten, die aufgrund ihrer Länge nicht nur eine breitere Materialbasis, sondern auch mehr Szenen ohne Geräusche und Musik bietet als ein herkömmlicher Kinofilm. Bei der vorliegenden Analyse war dies jedoch aufgrund ihrer spezifischen Fragestellung nicht möglich.

unterschiedlich zur Geltung (vgl. z.B. Blaseio 2003: 163). Dies liegt daran, dass der Ton beim Dreh eines Films meist<sup>46</sup> direkt eingespielt wird, während die Sprecher für die Synchronaufnahmen in einer schalldichten Kabine direkt vor einem Aufnahmemikrofon sitzen und Musik und Geräusche in Form des IT-Bandes später separat dazugemischt werden (vgl. Kapitel 2.4). Die Aufnahmetechnik unterscheidet sich in beiden Fällen somit u.U. erheblich, weshalb auf den Synchrononspuren die deutschen Stimmen i.d.R. mehr „in den Vordergrund“ treten.

Aus diesen Gründen waren für eine akustische und auditive Analyse nur solche Szenen des Films verwendbar, in welchen aus szenischen bzw. dramaturgischen Gründen die Stimme des Jokers weder durch Musik oder Soundeffekte, noch durch die Stimme eventueller Gesprächspartner überlagert ist. Davon gibt es im ganzen Film vier verschieden lange Szenen mit einer Gesamtlänge von rund 82 Sekunden (Video-Szene, Party-Szene, Vater-Szene, Gordon-Szene)<sup>47</sup>. Da dies bei einer Filmlänge von ca. 146 Minuten als Grundlage für eine aussagekräftige Analyse zu wenig schien, wurden zwei weitere Szenen mit längeren Äußerungspassagen des Jokers herangezogen (Batman-Szene und Hospital-Szene). In diesen sind zwar Musik und Hintergrundgeräusche vorhanden, sind jedoch so leise, dass zumindest eine rein auditive Analyse möglich sein sollte, um so in diesem Aspekt ein größeres Korpus zur Verfügung zu haben.

Wichtig bei der Auswahl der Szenen war außerdem, dass die Äußerungen des Jokers von anderen klar abgrenzbar sind und er sich, um eine größere Deutlichkeit zu gewährleisten, in normaler Sprechhaltung befindet. So wurde etwa das Ende der Batman-Szene, in welchem der Joker durch Batman drangsaliert wird und daher extrem gepresst spricht, nicht in die Analyse mit einbezogen (TC 1:25:28 ff.).

Um einer weiteren Bearbeitung zugänglich zu sein, wurden für jede Szene beide Tonspuren getrennt von der DVD herunter aufgezeichnet und als Sounddateien auf Festplatte abgespeichert. Dazu wurden mit Hilfe geeigneter Software die Szenen auf Deutsch und Englisch am PC abgespielt und der dargebotene Ton mit dem Soundbearbeitungsprogramm Audacity direkt über die Soundkarte aufgenommen. Hierzu stellt dieses Programm unter den Mikrofoneinstellungen die Funktion „Stereo-Mixer“ zur Verfügung, die automatisch die Soundkarte als Aufnahmequelle auswählt. So ist gewährleistet, dass die entsprechende Aufzeichnung die gleiche Qualität wie auf

---

<sup>46</sup> Gelegentlich werden auch bei sonst nicht synchronisierten Filmen Szenen nachsynchronisiert, wenn z.B. die Schauspieler aufgrund zu lauter Umgebungsgeräusche beim Dreh nicht zu verstehen waren (vgl. auch Blaseio 2003: 171, Anm. 2).

<sup>47</sup> Eine genaue inhaltliche Beschreibung der Szenen erfolgt in Kapitel 3.3.

DVD erhält. Anschließend wurden die Aufnahmen mit dem Programm Wave Editor der Firma Nero weiterbearbeitet. Die Äußerungen eventueller Dialogpartner des Jokers wurden herausgeschnitten, ebenso wurden die einzelnen Passagen so zurechtgeschnitten, dass sie in sich abgeschlossen und Szenenübergänge zu Beginn und am Ende, die „Reste“ anderer Szenen enthielten, nicht mehr zu hören waren. Wo die Aufnahme am PC zu leise erfolgt war, wurde dies korrigiert. Nun standen die einzelnen szenischen Äußerungen jeweils in deutscher und englischer Sprache als Wave-Dateien in CD-Qualität<sup>48</sup> einer Analyse zur Verfügung.

### 3.2.4 Analyse des Sprachmaterials

#### 3.2.4.1 Auditive Analyse

Verschiedentlich wurde schon ausgeführt, dass es sich beim Sprechschall um ein Komplexgebilde aus vielen verschiedenen Faktoren handelt. Das menschliche Hörorgan ist bestens in der Lage, diesem Komplexeindruck zahlreiche Informationen zu entnehmen. So hat sich in verschiedenen Untersuchungen gezeigt, dass das Ohr gerade im Bereich der sprachlichen Frequenzen äußerst sensibel reagiert und schon kleinste Veränderungen wahrnimmt (vgl. z.B. Eckert/Laver 1994: 161, Kranich 2002: 12). Aus diesem Grund hat sich auch in der Sprechwissenschaft – neben messtechnischen Verfahren – die auditive Analyse als Instrument zur Beschreibung der Merkmale des Sprechschalls bewährt (vgl. z.B. Heilmann 2002, Kranich 2002, Bose 2003, Redecker 2008).

Dabei handelt es sich um den innerlichen Nachvollzug eines Höreindrucks, indem einzelne Parameter durch geschulte Hörer gezielt abgehört werden. Gleichzeitig ermöglicht dies durch die Beschreibung der Ausprägung dieser Parameter, dass der hörbare Sprechprozess als Ganzes nachvollzogen werden kann (vgl. auch Eckert/Laver 1994: 5). Dazu bedienen sich geschulte Hörer der Methode des bottom-up-Hörens, das heißt, dass man von verschiedenen Merkmalen ausgehend zu einem Gesamteindruck gelangt, während Hören im Alltag auf umgekehrtem Wege funktioniert (vgl. Bose 2003: 91f.). Nur durch diese Vorgehensweise ist die ganzheitliche Analyse eines Höreindrucks möglich, während messtechnische Verfahren bestimmte Ausprägungen nur punktuell wiedergeben (vgl. Kranich 2002: 83).

---

<sup>48</sup> Bitrate 705 kBit/s, Abtastrate 44 kHz bei 16 Bit Abtastgröße.

Diese auditive Analyse des Sprachmaterials erfolgte auch in dieser Arbeit mit einem speziellen Analysebogen, worin die Ausprägung der in Kapitel 2.2.2 beschriebenen Parameter jeweils skalierten Werten zugeordnet wurde. Dabei orientiert sich die Einordnung und Skalierung der Merkmale an der Vorgehensweise in Braun/Heilmann (2006) und Bose (2003).

Die Ausprägung der Mehrzahl der Merkmale wurde dabei anhand einer bipolaren, fünfstufigen Intervallskala festgehalten, wobei die beiden Enden jeweils als klar voneinander abgrenzbare, verbale Antonyme den maximalen Ausprägungsbereich des Merkmals darstellen (so zum Beispiel „sehr leise – sehr laut“ für die Ausprägung der Lautstärke). Die Skalenmittelwerte stellen dabei „[s]oweit möglich [...] ‚neutrale‘ erwartete Wahrnehmungen für sachlich gesprochene deutsche [bzw. englische; C.H.] Äußerungen dar [...]“ (Bose 2003: 39). Diese Bezugsgrößen ergeben sich aus klaren auditiven, physiologisch beschreibbaren Eindrücken<sup>49</sup>. Für die einzelnen Phänomene des Stimmklangs wurde eine dreistufige einpolige Skala verwendet, da sich bei den verschiedenen Phonationsmodi keine negative Ausprägung, sondern nur ein „nicht bzw. mehr oder weniger vorhanden“ feststellen lässt.

Das in Kapitel 3.2.3 beschriebene Sprachmaterial wurde nun in der oben erläuterten Weise vom Verfasser abgehört und analysiert. Durch mehrmaliges Abhören wurde versucht, die so gewonnenen relativen und subjektiven Eindrücke weiter zu festigen (vgl. auch Bose 2003: 92). Dennoch merkt Bose (2003: 125) hier an, dass es nicht auszuschließen sei, „dass eine einzelne Person durch die außerordentlich intensive Beschäftigung mit dem Material aus einem bestimmten Untersuchungsinteresse heraus (unabsichtlich) geneigt sein könnte, sich bestimmte Bewertungen ‚zurechtzuhören‘“. Um diese allzu subjektive Färbung der Analyseergebnisse zu vermeiden, ist es unverzichtbar, das Material durch weniger involvierte, geschulte Kontrollhörer gegenhören zu lassen. Dabei erhöht sich der Grad der Objektivität, je mehr Hörer an der Analyse beteiligt waren. Aus diesem Grund wurde in der vorliegenden Untersuchung das Material zusätzlich durch zwei sprechwissenschaftlich bzw. logopädisch ausgebildete Kontrollhörerinnen getrennt analysiert, eine Zahl, die in der Mehrzahl der sprechwissenschaftlichen Arbeiten als ausreichend angesehen wird (vgl. etwa Kranich 2002, Heilmann 2002). Anschließend wurden gemeinsam die gewonnenen Ergebnisse verglichen und auf ihr mögliches Zustandekommen hin diskutiert, um eine nachvollziehbare Eindrucksbeschreibung zu gewährleisten.

---

<sup>49</sup> Vgl. etwa die Beschreibungen für die „Settings“ des Stimmklangs in Kapitel 2.2.2.

### 3.2.4.2 Akustische Analyse

Um die durch die auditive Analyse gewonnenen Ergebnisse zu untermauern bzw. zu objektivieren, wurden die passenden Szenen (vgl. Kapitel 3.2.3) zusätzlich einer messphonetischen, akustischen Analyse unterzogen (vgl. hierzu etwa Bose 2003: 93). Denn im Gegensatz zum menschlichen Gehör, das nur eine relative Einschätzung der Merkmale ermöglicht<sup>50</sup>, sind Computerprogramme in der Lage, klare, absolut miteinander vergleichbare Messwerte zu liefern. Die auditive Einschätzung der Tonhöhe beispielsweise ist wie alle Eindrücke subjektiv etwa an vorhergehende Erfahrungen (vgl. Kapitel 2.2.1) oder je nach sprachlicher Umgebung an parallele Eindrücke wie Sprechspannung und Pausen geknüpft (vgl. Heilmann 2004: 108). Um hier solche Phänomene nachvollziehen und ausschließen zu können, sind entsprechende Messwerte nötig. Deshalb wird die Tonhöhe durch die Messung der Grundfrequenz als Richtwert überprüft. Auch das Tonhöhenintervall, der sogenannte Range, wird durch den tiefsten und höchsten  $f_0$ -Wert einer Ausdruckseinheit mess- und damit vergleichbar. Ebenso kann der Tonhöhenverlauf einer sprachlichen Äußerung angezeigt werden. So lassen sich heute schon mit einfach zugänglichen Programmen entsprechende Daten gewinnen, welche für die erwähnten Zusammenhänge wertvolle Informationen liefern. Als Programm wurde hier das frei verfügbare Praat<sup>51</sup> verwendet. Die Sprachdateien wurden nacheinander eingelesen und die entsprechenden Werte notiert bzw. die Melodieverlaufskurven als Graphiken abgespeichert.

Dennoch ist es bei einer Analyse wie der vorliegenden wichtig, die entsprechenden akustischen Ergebnisse immer zu den auditiv gewonnenen Einschätzungen in Beziehung zu setzen. Dies hat im wesentlichen zwei Gründe: Zum einen können auch Computermessungen fehlerhaft sein. Zum anderen besteht zwischen auditiven und akustischen Merkmalen keine Korrelation. Im Gegenteil, ein auditives Merkmal wird oft durch mehrere akustische Parameter konstituiert bzw. kann diesen entsprechen (vgl. Kehrein 2002: 99f.). Deshalb darf ein akustischer Parameter nie als alleinige Vergleichsgrundlage herangezogen werden (vgl. Kranich 2002: 83, Bose 2003: 93). Besonders deutlich wird dies ebenfalls am Beispiel der Korrelation von Grundfrequenz und Tonhöhe: Wie schon in Kapitel 2.2.2 ausgeführt, stellt die auditiv wahrnehmbare

---

<sup>50</sup> von seltenen Ausnahmen wie der Fähigkeit zum „absoluten Gehör“ einmal abgesehen.

<sup>51</sup> Weitere Informationen unter <http://www.fon.hum.uva.nl/praat>. Verwendet wurde Version 5.1.04.



Tonhöhe einen Komplex aus mehreren messbaren Teilschwingungen sowie anderer Größen wie der Intensität und Dauer dar, von deren Ausprägung die Wahrnehmung der Tonhöhe ebenfalls abhängig ist (vgl. auch Niebuhr 2007: 5). Deshalb wird sie durch die messbare Grundfrequenz nur zum Teil wiedergegeben und darf vor allem bei der Interpretation der Ergebnisse nicht mit ihr gleichgesetzt werden, wie dies in einigen Untersuchungen der Fall ist<sup>52</sup>. Dennoch erfassen beide Methoden unterschiedliche Bereiche, wodurch sich manche Nachteile der einen durch die jeweils andere vermindern bzw. ausgleichen lassen, was für ein kombiniertes Vorgehen in der Analyse spricht (vgl. etwa auch Kranich 2002: 84). Auch Heilmann (2004: 109) stellt als Fazit einer komplexen Untersuchung fest, dass „[b]eide Leistungen [...] für sprechwissenschaftliche und phonetische Untersuchungen und die Wahrnehmung von Sprechrealität unabdingbar“ seien. Aus diesem Grund kommen auch in dieser Arbeit beide Verfahren zur Anwendung.

#### *3.2.4.3 Analyse der temporalen Parameter*

Zusätzlich zu diesen Daten wurden instrumentell-perzeptiv die temporalen Parameter jeder Äußerungssequenz erfasst. Denn verschiedene Untersuchungen haben gezeigt, dass diese in der Sprechwirkung ebenfalls eine Rolle spielen (vgl. etwa Bartels/Sendlmeier 1997, Kowal 1991). Dazu zählen Pausenzahl und –dauer, sowie Artikulations- und Silbenrate.

Um diese Werte zuverlässig bestimmen zu können, ist eine klare Definition des Begriffs „Pause“ unerlässlich. Diese Arbeit folgt dabei der Begriffsbestimmung von Kowal (1991: 34), wonach Pausen diejenigen Intervalle einer Äußerung bezeichnen, „in denen für eine perzeptuell oder instrumentell erfasste Mindestdauer keine Artikulation einer bestimmten Mindestintensität zu beobachten ist und die durch die Artikulation desselben Sprechers oder verschiedener Sprecher begrenzt sind [...]“. Es handelt sich hierbei somit um eine akustisch-auditive Definition und keine rein artikulatorische, die etwa nicht redeeinleitende Artikulationsbewegungen ohne hörbare Lautproduktion mit als Pausen klassifizierte. Denn diese sind bei reinen Sprachaufnahmen, wie sie die vorliegende Analyse zum Gegenstand hat, nicht zu erfassen. Bei der Festlegung der erwähnten Mindestdauer ist zu beachten, dass das entsprechende Intervall nicht so kurz gewählt wird, dass auch Artikulationspausen innerhalb einzelner Silben mit erfasst

---

<sup>52</sup> vgl. zu diesem Aspekt Kehrein (2002: 80).

werden, die etwa notwendigerweise bei der Artikulation von Verschlusslauten auftreten. Bisherige Untersuchungen haben außerdem gezeigt, dass daneben Pausen zwischen 0,13 und 0,27 Sekunden für die Sprachwahrnehmung bedeutsam sein können (vgl. Kowal 1991: 35), deshalb werden alle Pausen mit einer Mindestdauer von 0,1 Sekunden in die Analyse mit einbezogen. Es handelt sich hierbei somit um klar wahrnehmbare Intervalle ohne Artikulation zwischen Wörtern.

Dabei genügt für die vorliegende Arbeit eine Makroanalyse der Pausenstruktur mit Angaben zur Höhe des Anteils der Pausenzeit an der gesamten Sprechzeit und der Pausenhäufigkeit (vgl. Kowal 1991: 42f.). Da zwei verschiedene Sprachen betrachtet und Pausen lediglich als temporales Merkmal untersucht werden sollen, erübrigt sich eine funktionale und verteilungsabhängige Beschreibung.

Da sich eine rein auditive Pausenerfassung als problematisch herausgestellt hat (vgl. die Beispiele in Kowal 1991: 37), wird in dieser Analyse ein instrumentell-perzeptiver Ansatz verwendet (vgl. Kowal 1991: 39). Das heißt, mit Hilfe des Computerprogramms Nero Wave Editor<sup>®</sup>, das in der Lage ist, den Intensitätsverlauf einer Äußerung in Abhängigkeit zur Zeit darzustellen, wurden Pausendauer und -häufigkeit für jede Szene anhand des dargestellten Oszillographen sowie durch paralleles Hören manuell bestimmt. Die Reliabilität dieses halb instrumentellen Verfahrens bezeichnet Kowal (1991: 39) als „im allgemeinen sehr hoch“. Es eignet sich außerdem insbesondere für nicht unter Laborbedingungen erstellte Aufnahmen wie in der vorliegenden Untersuchung, „da der (trainierte) Auswerter als Kontrollinstanz eingeschaltet ist, um gegebenenfalls zwischen Signal und Geräusch zu unterscheiden“ (Kowal 1991: 39). Dies war insbesondere bei den zwei rein auditiv analysierbaren Szenen von Bedeutung, da diese mit rein apparativen Methoden sonst nicht in eine temporale Analyse hätten mit einbezogen werden können. Daneben hat dieses Verfahren den „Vorteil, daß es eine intensive Beschäftigung mit den Sprachdaten erfordert, ohne die Einschränkungen des rein perzeptiven Verfahrens zu teilen“ (Kowal 1991: 39). Es stellt somit in den in diesem Kapitel erläuterten Analysemethoden eine Zwischenstufe zu auditiven bzw. akustischen Vorgehensweisen dar. Hinzu kommt noch, dass die Entwicklung digitaler oszillographischer Aufzeichnungsprogramme in den letzten Jahren erhebliche Fortschritte gemacht hat.

Als Maß für die Sprechgeschwindigkeit wurde in dieser Arbeit die Artikulationsrate (AR) ermittelt, was die durchschnittliche Anzahl der artikulierten Silben pro Sekunde

abzüglich der Sprechpausen bezeichnet<sup>53</sup>. Diese gibt den zeitlichen Verlauf eines reinen<sup>54</sup> Sprechereignisses recht zuverlässig wieder, wobei die folgenden methodischen Faktoren berücksichtigt wurden:

1. Als Sprecheinheit wird die Silbe betrachtet, da sie „reliabel identifiziert werden kann und [...] in verschiedenen Sprachen nur geringfügig variiert“ (Kowal 1991: 95).
2. Pausen sind in ihrer Länge klar klassifiziert (vgl. oben), um nicht durch verschieden eingerechnete Pausenlängen keine vergleichbaren Werte für die verschiedenen Szenen zu erhalten. Dadurch, dass Geräusche und Äußerungen anderer Gesprächspartner aus den Szenen herausgeschnitten wurden, wurden nur die Pausen Heath Ledgers zwischen seinen Äußerungen mit einbezogen.
3. Als „Äußerung“ werden hier alle hörbar artikulierten Einheiten aufgefasst, also auch Korrekturerscheinungen und Füllwörter, was nach Kowal unerlässlich ist, um den genauen „Prozeß der schrittweisen Formulierung der Gedanken“ (1991: 97; Hervorh. von der Autorin) als Bestandteil des Sprechvorgangs in der Analyse mit zu berücksichtigen.
4. Die Angaben zur AR beziehen sich immer auf die gesamte Szene und nicht nur auf einzelne Phrasen, da es hier um den globalen Vergleich und nicht so sehr um einzelne Schwankungen der AR innerhalb einer Äußerung gehen soll.

### 3.3 Ergebnisse

In diesem Kapitel werden nun die Ergebnisse der auditiven und akustischen Analyse für die englische und deutsche Fassung der verwendeten Szenen referiert<sup>55</sup>. Dabei werden zunächst die jeweiligen Ausprägungen der einzelnen Parameter im szenischen Kontext verglichen und miteinander in Bezug gesetzt. Kapitel 3.3.7 fasst dann die generalisierbaren, szenenübergreifenden Ergebnisse der Analyse für beide Stimmen zusammen und versucht, mögliche Zusammenhänge der Parameter aufzuzeigen.

---

<sup>53</sup> Kowal (1991: 93) spricht hier von „Artikulationsgeschwindigkeit“. Um jedoch die Art der Angabe zu verdeutlichen, wird der Terminus „Rate“ beibehalten (vgl. auch Bartels/Sendlmeier 1997).

<sup>54</sup> Weil nur die Zeiteinheiten erfasst werden, in welchen auch tatsächlich Sprache zu hören ist (vgl. Kowal 1991: 93).

<sup>55</sup> Alle Werte der akustischen Messung finden sich in der Übersichtstabelle auf Seite 82 im Anhang.

### 3.3.1 Video-Szene

In dieser Szene ist in Bruce Waynes Fernseher eine Videobotschaft zu sehen und zu hören, in welcher der Joker damit droht, jeden Tag Menschen zu ermorden, sollte Batman nicht seine wahre Identität preisgeben. Dazu hält der Joker eine Geisel gefangen, die sich als Batman verkleidet hat. Im analysierten Teil der Szene (TC 41:22 bis 41:48) wendet sich der Joker direkt an den Fernsehzuschauer. Anschließend tötet er die Geisel. Der erste Teil der Szene, in welchem sich der Joker in Dialog mit der Geisel befindet, wurde aufgrund von Äußerungsüberschneidungen nicht in die Analyse mit einbezogen.

#### *Temporale Merkmale:*

Auffällig ist, dass hier die deutsche Stimme auditiv leicht schneller wahrgenommen wird als die englische. Die Artikulationsrate (AR) ist mit 3,81 Silben pro Sekunde im Deutschen gegenüber 3,29 Silben pro Sekunde im Englischen jedoch nur ca. eine halbe Silbe schneller. Ein größerer Unterschied ergibt sich bei der Betrachtung von Pausenzeit und Sprechzeit: Der Synchronsprecher Simon Jäger benötigt ca. 1,4 Sekunden mehr Sprechzeit als Heath Ledger, denn er muss 14 Silben mehr artikulieren. Trotzdem darf er, damit die Synchronität zwischen Bild und Ton gewahrt bleibt, die Gesamtlänge der Szene nicht überschreiten. Dies erreicht er, indem er insgesamt ca. 1,6 Sekunden weniger Pausen macht. Die einzelnen Äußerungen folgen bei annähernd gleicher AR in derselben Zeit im Deutschen somit schneller aufeinander, was vermutlich als Erhöhung des Sprechtempos wahrgenommen wird (vgl. Zilliken 1991: 34).

#### *Dynamische Merkmale:*

Beide Stimmen wurden als lauter, mit weiterem Lautstärkebereich und leicht stärkerer Lautstärkevariation wahrgenommen, somit konnte kein klarer Unterschied zwischen beiden Stimmen festgestellt werden. Auch die Intensitätswerte unterscheiden sich mit im Durchschnitt 2 dB nur geringfügig.

#### *Melodische Merkmale:*

Bezüglich der Tonhöhe werden beide Sprecher auditiv als hoch beurteilt, mit weitem bis sehr weitem Tonhöhenbereich und weiter bis sehr weiter Tonhöhenvariation. Diese Einschätzung wird durch die akustische Analyse bestätigt: Der englische Sprecher hat mit 147,9 Hz eine recht hohe Grundfrequenz, die der deutsche mit 190,8 Hz noch einmal deutlich übertrifft. Auch der Range ist mit 368,43 im Englischen bzw. 353,33 im Deutschen für beide Stimmen der größte unter den analysierten Szenen, Heath Ledger

befindet sich zeitweilig sogar maximal auf 416,1 Hz, Simon Jäger knapp bei 400 Hz. Dies mag damit zusammenhängen, dass der Joker hier mit sehr starken melodischen

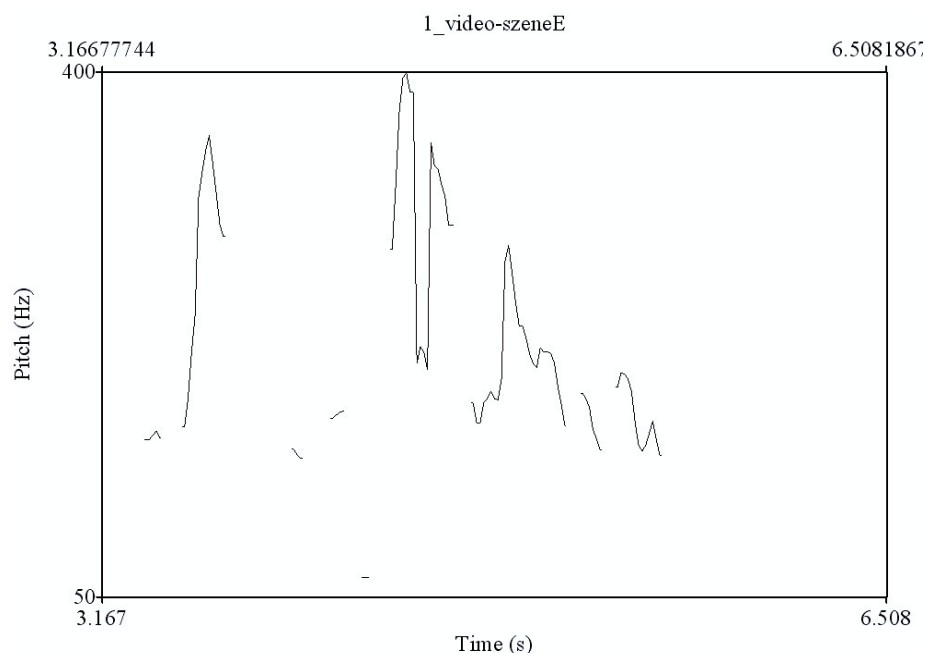


Abb. 3.2: Grundfrequenzverlauf des Satzes:  
*You see, this is how crazy Batman's made Gotham*

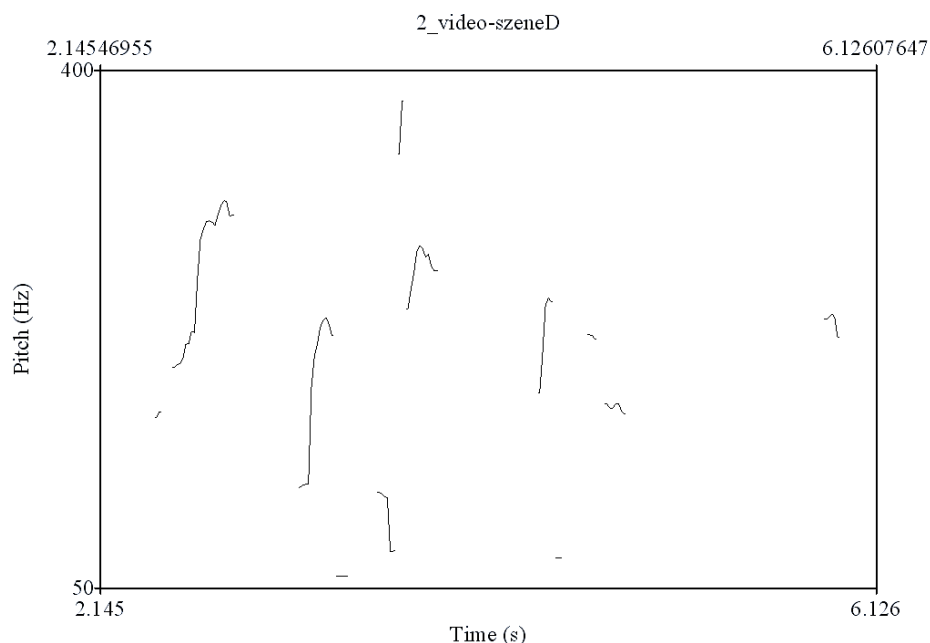


Abb. 3.3: Grundfrequenzverlauf des Satzes:  
*Seht ihr, so verrückt hat Batman Gotham gemacht*

Akzenten spricht, um  
wortwörtlich seine Verrücktheit

zu unterstreichen. In den obigen Abbildungen ist dies an einer Äußerung am Anfang der Sequenz illustriert, wo der Joker seine Verrücktheit explizit thematisiert. Beide Sprecher nutzen eine große Tonhöhenvariation mit dem Gipfel von  $F_0$  auf *crazy* bzw. *verrückt*, wobei der deutsche (Abb. 3.4) deutlich mehr Sprünge und weniger gleitende Melodiebewegungen einsetzt. Auch spricht er gegen Ende der Äußerung mit sehr großer

Sprechspannung und sehr gepresst, weshalb kaum noch eine messbare Schwingung vorhanden ist.

*Artikulatorische Merkmale:*

Hier wird Heath Ledgers Stimme als schärfer artikulierend wahrgenommen, er macht hörbare „Schmatzgeräusche“ mit den Lippen.

*Stimmklang:*

Hier zeigen sich kaum Auffälligkeiten bis auf eine leicht behauchte und knarrende Anregung in der deutschen Stimme. Die Wahrnehmung des weiten Range bewirkt eine zeitweise Einordnung beider Stimmen in das Falsettregister, jedoch benutzen beide Sprecher auch den modal voice. Hinsichtlich Klangfülle werden beide Stimmen klangärmer als normal eingeschätzt, die Klangfarbe der deutschen Stimme jedoch als heller und die englische als dunkler.

*Sprechspannung:*

Beide Stimmen befinden sich in allen Spannungen etwa eine Stufe über dem normalen Level, wobei die deutsche Stimme vor allem laryngal sehr viel Spannung aufweist.

### 3.3.2 Party-Szene

In dieser Szene (TC 47:13 ff.) stürmt der Joker eine Spendengala, die Bruce Wayne zu Ehren des neuen Staatsanwalts Harvey Dent initiiert hat. Nachdem der Joker bereits Polizeichef Lobe und eine Richterin getötet hat, will er nun im Auftrag der Mafia Dent entführen. Er bedroht die Gäste, um zu erfahren, wo sich Harvey Dent versteckt hält. Wegen einer die Sprache überlagernde Geräuschkulisse und der ab TC 47:42 einsetzenden Musik konnte hier wieder nur ein Teil der Szene (TC 47:13 bis 47:42) analysiert werden. Dennoch wurden die möglichen Äußerungen vor allem deshalb mit in die Analyse aufgenommen, weil es sich hierbei um jene Szene handelt, die den Nutzern auf YouTube zur Verfügung stand und sich entsprechende Kommentare teilweise direkt darauf beziehen (vgl. Kapitel 3.2.2).

*Temporale Merkmale:*

Auditiv war hier bezüglich des Sprechtempos kein Unterschied zwischen beiden Stimmen festzustellen. Bei Betrachtung der Artikulationsrate fällt auf, dass Heath Ledger mit 4,29 Silben/Sekunde etwas schneller artikuliert als der deutsche Sprecher (3,98 Silben/Sekunde), was im Vergleich zu den restlichen Szenen eine klare Ausnahme darstellt, diesen also zu widersprechen scheint. Kein gegensätzliches Bild ergibt sich

dabei jedoch erneut beim Vergleich der Pausenzeit: Simon Jäger macht 0,9 Sekunden weniger Sprechpausen als Heath Ledger, was mit 45% Unterschied im Vergleich die höchste Differenz ausmacht. Es scheint also, als ob der deutsche Sprecher wieder eher die Pausen zur Artikulation nutzt, als generell schneller zu sprechen. Warum er dabei jedoch im Schnitt langsamer artikuliert als sein englischer Gegenpart, lässt sich vielleicht dadurch erklären, dass er diesmal nur vier Silben mehr unterbringen muss, was im Vergleich mit den anderen Szenen die niedrigste Differenz an Silben (13,8%) darstellt.

#### *Dynamische Merkmale:*

Hinsichtlich Lautstärke und dynamische Akzente sind beide Stimmen auch hier sehr aneinander angeglichen, was sich durch die kaum unterschiedlichen Intensitätswerte bestätigt.

#### *Melodische Merkmale:*

Auditiv werden Tonhöhe, Tonhöhenbereich und Tonhöhenvariation der englischen Stimme als nicht auffällig beurteilt. Die deutsche Stimme wird jedoch mit einem leicht erweiterten Tonhöhenbereich und größerer Tonhöhenvariation wahrgenommen. Betrachtet man die akustischen Messwerte, so fällt auf, dass beide Stimmen mit 178,6 Hz (englisch) zu 184,7 Hz (deutsch) tatsächlich ebenfalls eine ähnliche, wenn auch recht hohe  $F_0$  aufweisen. Die englische Stimme ist auch vom Range her recht hoch

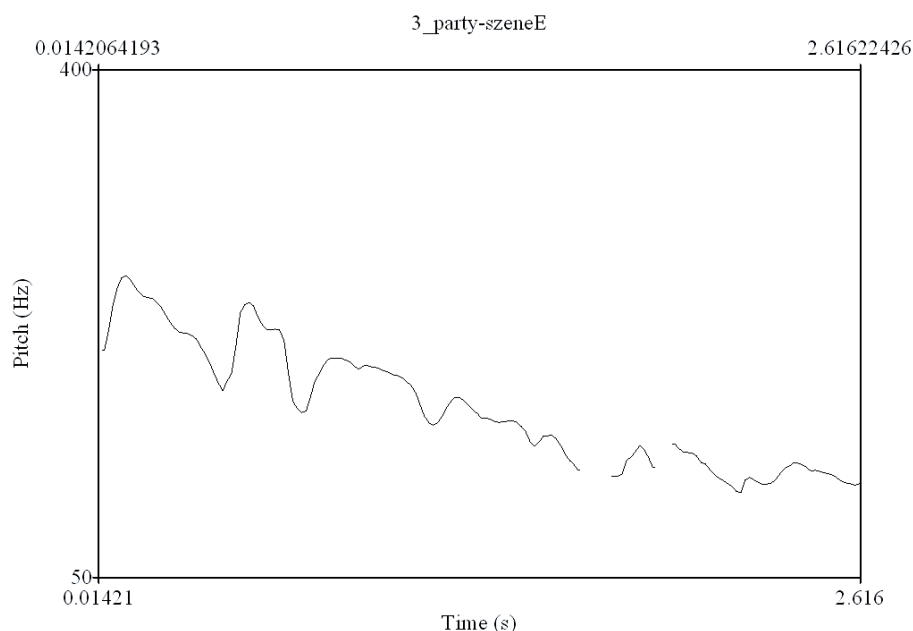


Abb. 3.4: Grundfrequenzverlauf des Satzes: *Good evening Ladies and Gentlemen*

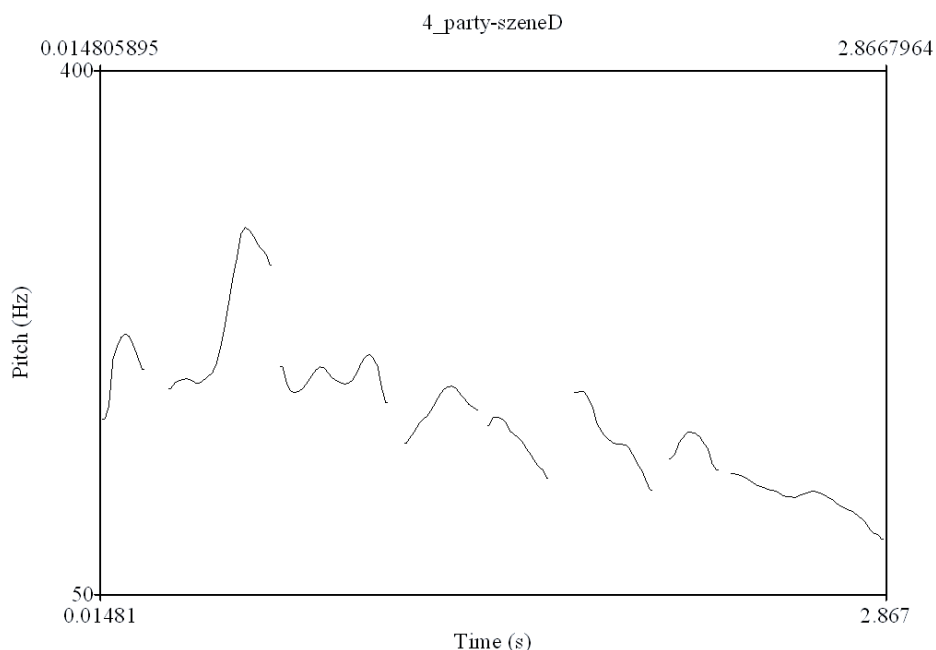


Abb. 3.5: Grundfrequenzverlauf des Satzes: *Guten Abend, Ladys und Gentlemen*

angelegt, mit 101,1 Hz als niedrigstem Wert, der Gesamtumfang ist mit 176,9 jedoch recht eng. Der deutsche Sprecher nimmt mit 232,3 zwar auch seinen zweitgeringsten Range ein, hat aber dennoch einen größeren Umfang als der englische, vor allem weil er zeitweise mit einer  $F_0$  unter 100 Hz spricht. Daneben weist auch der Grundfrequenzverlauf eine größere Variation auf. Dies sei am Beispiel des Eingangssatzes der Szene verdeutlicht. Dieser lässt sich besonders gut vergleichen, weil er in beiden Sprachen zu einem Großteil aus ähnlichen Phonen besteht (vgl. Abb. 3.5 u. 3.6). Hier sieht man deutlich, dass beide Sprecher einen global fallenden Melodieverlauf nutzen. Der deutsche Sprecher macht jedoch erneut mehr Sprünge, die zudem besonders im ersten Teil des Satzes deutlich größer ausfallen als Ledgers Variationen.

#### *Artikulatorische Merkmale:*

Heath Ledger isst in Teilen der Szene ein paar Partyhäppchen, wodurch sich die Artikulation im Ansatzrohr leicht nach hinten verlagert. Dies wird vom deutschen Sprecher nahezu gleich übernommen, sodass eine Kongruenz zwischen Bild und Ton gewährleistet ist.

#### *Stimmklang:*

Heath Ledger zeigt hier bezüglich des Stimmklangs keine Auffälligkeiten, der deutsche Sprecher besitzt eine leicht knarrende Anregung. Auffällig sind die wahrgenommenen Unterschiede bezüglich Klangfarbe und -fülle: Die englische Stimme wird als klangvoll und dunkel, die deutsche Stimme trotz der faukalen Weite wegen des „Essens“ aber als im Vergleich weniger klangvoll und deutlich heller wahrgenommen.



*Sprechspannung:*

Beide Stimmen besitzen durch fehlende velopharyngale Spannung Nasalität, wobei die deutsche im Gegensatz zur englischen als sehr nasal wahrgenommen wird. Beide Sprecher sind supralaryngal leicht gespannt, während die laryngale Spannung durch die große faukale Weite am Schluss sehr gering ist. Insgesamt sind beide Sprecher also leicht gespannt, mit höherer Nasalität im Deutschen.

### 3.3.3 Vater-Szene

In dieser sehr kurzen Sequenz (knapp fünf Sekunden) spricht der Joker auf der oben erwähnten Party einen Gast direkt an, nachdem dieser ihm erklärt hat: „We are not intimidated by thugs“ (Deutsch: „Wir lassen uns von Strolchen wie Ihnen nicht einschüchtern“). Der Joker teilt dem Mann mit, dass er ihn an seinen Vater erinnere und dass er diesen gehasst habe, wobei er bei der letzten Äußerung den Mann am Kragen packt und ihm ein Messer an die Kehle hält (TC 48:14 bis 48:22, das Abstellen eines Glases wurde aus der Sequenz herausgeschnitten). Diese Szene befindet sich am Ende des YouTube-Videos und die Kommentare beziehen sich auch hier teilweise direkt darauf.

*Temporale Merkmale:*

Auditiv wird bei beiden Stimmen die Sprechgeschwindigkeit als normal wahrgenommen. Dies verwundert jedoch, denn dem deutschen Sprecher bleibt in dieser kurzen Sequenz, wo er ebenfalls fünf Silben mehr zu sprechen hat, kaum Zeit für Pausen und er artikuliert daher im Schnitt eine ganze Silbe schneller. Dass dies auditiv nicht wahrgenommen wird, spricht wohl dafür, dass die Wahrnehmung der Sprechgeschwindigkeit nicht allein über die AR erfolgt. Allerdings ist diese Szene auch sehr kurz, sodass wenig zuverlässige Aussagen über ein sehr variables Phänomen wie die Sprechgeschwindigkeit getroffen werden können.

*Dynamische Merkmale:*

Auch in dieser Szene gab es bezüglich dieser Parameter keine Auffälligkeiten.

*Melodische Merkmale:*

Tonhöhe und Tonhöhenbereich werden für beide Stimmen ähnlich wahrgenommen, während in der deutschen Stimme mehr Tonvariation zu hören ist. Auffällig ist die extrem gepresste Sprechweise im zweiten Teil der Äußerung, in welchem die deutsche Stimme regelrecht diplophon klingt. Das bedeutet, der deutsche Sprecher spricht mit so

hoher Spannung und so großem Atemdruck an den Stimmlippen, dass diese sehr unregelmäßig schwingen und zwei Grundfrequenzen sich überlagern. Das Analyseprogramm konnte für diesen Teil der deutschen Äußerung auch keinen Frequenzverlauf mehr bestimmen. Deshalb wurde nur der erste Teil der Szene für beide Stimmen analysiert. Die englische Stimme geht an dieser Stelle bis 50 Hz, bleibt also im für sie normalen Bereich, während sich die deutsche Stimme hier in allen Parametern sehr vom ersten Teil unterscheidet<sup>56</sup>.

Bezüglich der Grundfrequenz ist Heath Ledger im ersten Teil der Äußerung mit 181,1 Hz im Vergleich für ihn recht hoch jedoch sehr nah an der deutschen Stimme (190,7 Hz), was deren auditive Wahrnehmung als etwa gleich hoch bestätigt. Dies zeigt, dass die auditive Einschätzung der Stimmhöhe auch von äußeren Faktoren wie im selben Kontext gehörten Stimmen abhängig ist. Auch bezüglich Range sind sie sich sehr ähnlich, wobei die deutsche Stimme mit 188,01 erneut den größeren Umfang hat.

*Artikulatorische Merkmale:*

Auch in dieser Szene zeigen sich in beiden Sprachen keine Unterschiede, beide sprechen die erste Äußerung mit recht weiter Kieferöffnung, im zweiten Teil ist der Kiefer extrem geschlossen, bei hoher Artikulationsspannung.

*Stimmklang:*

Beide Stimmen haben eine leicht behauchte Anregung im ersten Teil der Äußerung und werden anschließend als rauh wahrgenommen. Unterschiede gibt es auch in dieser Szene bezüglich der Klangfarbe: Die deutsche Stimme wird insgesamt erneut als heller beurteilt, die englische dagegen dunkler.

*Sprechspannung:*

Hier ist die englische Stimme nasalere als die deutsche, während sonst die deutsche Stimme im Vergleich insgesamt deutlich mehr Sprechspannung aufweist.

### 3.3.4 Gordon-Szene

Diese Szene setzt unmittelbar nach der Festnahme des Jokers durch Commissioner Gordon an. Die Polizei hat inzwischen festgestellt, dass Harvey Dent von den Leuten des Jokers entführt wurde, und der Joker wird nun von Gordon dazu verhört, wobei er nichts zugibt, aber klar mit seinem Wissen um die Machtlosigkeit der Polizei spielt (TC 1:22:18 bis 1:23:36).

---

<sup>56</sup> Nutzer PrincessSalira fasst dies als Eindruck so zusammen: „Wow, der letzte Satz kommt im Deutschen fast noch verbitterter rüber als im Original!“.

*Temporale Merkmale:*

Diese Szene stellt mit knapp 50 Sekunden die längste unter den akustisch analysierten Szenen dar. Deshalb lässt sich hier die Artikulationsrate erstmals für einen längeren Zeitraum feststellen. Mit 3,47 (englisch) zu 3,43 Silben/Sekunde (deutsch) ist die AR für beide Sprachen hier eher langsam, was die auditive Wahrnehmung bestätigt. Für den deutschen Sprecher ist es der niedrigste Wert in den analysierten Szenen, für den englischen der zweitniedrigste. Das entspricht dem aufreizend ruhigen Sprechen des Jokers angesichts seiner bewussten Überlegenheit, das hier wahrgenommen wird. Dennoch hat der deutsche Sprecher wieder 18 Silben mehr zu artikulieren, was er erneut durch eine Verkürzung der Pausenzeit erreicht. Das heißt, er ist gezwungen, die einzelnen Phrasen der langen Äußerungspassagen, wenn auch nur minimal, doch schneller aufeinander folgen zu lassen, als Heath Ledger das in seiner Vorlage tut.

*Dynamische Merkmale:*

Die Lautstärke in beiden Sprachen wird als deutlich niedriger wahrgenommen, was die Intensitätswerte bestätigen. Der deutsche Sprecher variiert die Lautstärke weiter, wobei der englische mehr dynamische Akzente setzt.

*Melodische Merkmale:*

Hier zeigen sich zwischen beiden Stimmen recht deutliche Unterschiede. Die Tonhöhe des deutschen Sprechers wurde im Gegensatz zum englischen eher mit hoch beurteilt.

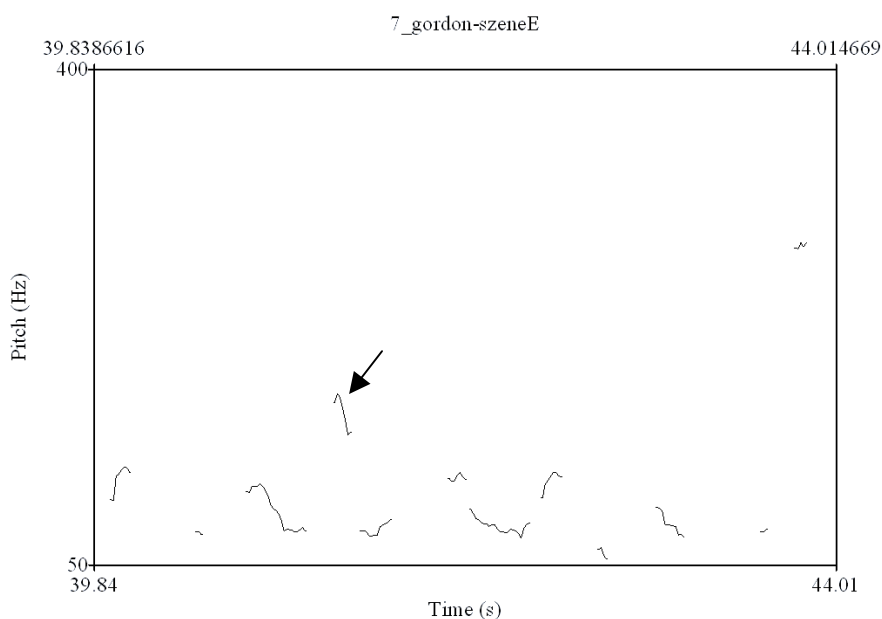
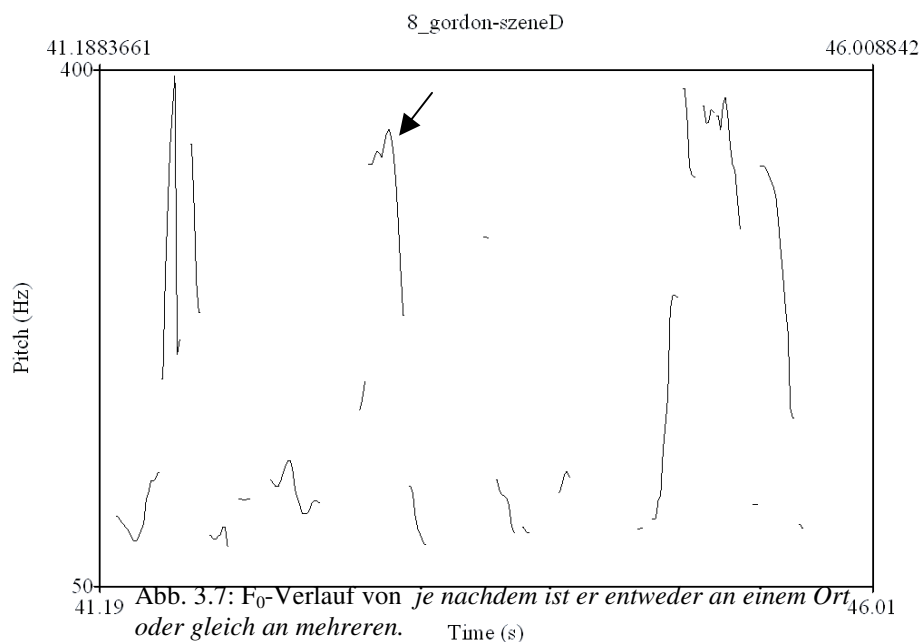


Abb. 3.6: F<sub>0</sub>-Verlauf von *depending on the time, he may be on one spot or several.*

Die Tonhöhenvariation Heath Ledgers ist eher niedrig, bei kleinerem Tonhöhenbereich. Der deutsche Sprecher weist dagegen sehr viel wahrnehmbare Tonhöhenvariation auf. Diese Eindrücke werden durch die akustischen Messwerte bestätigt. Mit 128,1 Hz hat

die deutsche Stimme hier eine eher niedrige Grundfrequenz bezogen auf die sonstigen Werte, die jedoch im Vergleich zum englischen (106 Hz) deutlich höher liegt, was zeigt,



dass die Wahrnehmung der Tonhöhe bezogen auf die Grundfrequenz relativ ist und etwa von anderen Stimmen abhängt. Auch bezüglich Range schöpft der deutsche Sprecher mit 349,8 zwar nicht das Maximum aus, im Vergleich ist der des englischen mit 179,6 jedoch deutlich kleiner, zumal er hier seinen zweitniedrigsten, der deutsche Sprecher jedoch den zweithöchsten Wert hat. Obwohl es sich inhaltlich um die gleiche Szene handelt, schöpft der deutsche Sprecher seinen Range hier also deutlich mehr aus. Er erreicht mit 397,2 Hz außerdem sein zweitgrößtes F<sub>0</sub>-Maximum, während Heath Ledger im Vergleich mit 227,6 Hz sein kleinstes Maximum erhält. Hier treten also – vielleicht ob der Szenenlänge – recht deutliche Unterschiede beider Sprecher in der Melodisierung zu Tage. In der obigen Äußerung gegen Ende der Szene ist dies recht deutlich sichtbar (Die Antwort des Jokers auf die Frage nach dem Aufenthaltsort von Harvey Dent). Der Synchronsprecher legt gleich zu Beginn einen sehr deutlichen melodischen Akzent auf *nachdém*, wobei er fast sein F<sub>0</sub>-Maximum erreicht, während Heath Ledgers Akzent auf *tíme* sichtbar niedriger ausfällt, aber immer noch gut zu erkennen ist (Pfeile). Auch im weiteren Verlauf der Äußerung zeigen sich wieder deutliche Melodiesprünge in der deutschen Äußerung (*oder gleich an méhreren*), während Heath Ledger im letzten Teil (*or sév'ral*) sehr rau klingt und kaum noch messbare Schwingungen auftreten, er somit eher mit dynamischen Akzenten arbeitet.

*Artikulatorische Merkmale:*

Artikulatorisch zeigt Heath Ledger eine höhere Variation in der Spannung der Muskulatur.

*Stimmklang:*

Beide Stimmen haben eine sehr knarrende Anregung, die deutsche ist daneben noch behaucht und zeigt sogar zeitweilig Flüsteranteil, der Sprecher nimmt die Stimme also ganz zurück. Daneben zeigt er auch über die ganze Szene hier mehr Variationen, auch bezüglich Klangfülle und -farbe, so dass ein einheitliches Wahrnehmungsbild schwierig ist. Trotzdem wird die deutsche Stimme erneut als heller beurteilt. Die englische Stimme wird dagegen sehr einheitlich als klangvoller und dunkler wahrgenommen.

*Sprechspannung:*

Erneut wird die deutsche Stimme als nasal wahrgenommen, jedoch ist der Ausprägungsgrad schwer zu bestimmen. Beide Stimmen werden jedoch einheitlich insgesamt als unterspannt beurteilt, was der sehr knarrenden Anregung aufgrund geringer laryngaler Spannung entspricht. Heath Ledger zeigt jedoch bezüglich der supralaryngalen Spannung mehr Variation, zum Beispiel bei der Phrase *or several* gegen Ende der Szene.

### 3.3.5 Batman-Szene

Diese Szene (TC 1:24:02 bis 1:25:28) schließt, bezogen auf den Handlungsablauf, unmittelbar an die Gordon-Szene an. Nachdem Gordon im Verhör nicht die gewünschten Informationen erhalten hat, verlässt er den Verhörraum. Plötzlich wird das Licht eingeschaltet und man sieht, dass Batman die ganze Zeit über hinter dem Joker stand. Er will nun seinerseits durch eine etwas härtere Gangart den Joker zum Reden bringen. Dieser teilt Batman auch bereitwillig und nahezu genüsslich seine Motivation mit: Seine Verbrechen dienen allein dazu, durch die Verbreitung von Panik und Chaos die Moral der Leute zu untergraben und zu beweisen, dass jeder Mensch im Grunde unmoralisch ist, wenn ihm die äußeren Umstände nur die Gelegenheit dazu geben. Dabei wurde der Anfang der Szene (1:23:40 bis 1:24:00) wegen zu vielen Hintergrundgeräuschen wie Licht einschalten, Stühle rücken etc. nicht in die Analyse mit einbezogen. Außerdem konnte diese Szene wie auch die folgende wegen der leise einsetzenden Hintergrundmusik nur auditiv analysiert werden.

### *Temporale Merkmale:*

Die deutschen Äußerungen werden wieder als leicht schneller in der Sprechgeschwindigkeit wahrgenommen. Dies bestätigt eine genauere Analyse: Der Synchronsprecher macht 1,2 Sekunden weniger Pausen und artikuliert mit einer AR von 4,3 Silben/Sekunde auch im Durchschnitt 0,7 Silben mehr pro Sekunde als Heath Ledger, da er 46 Silben mehr in der gleichen Zeit unterbringen muss. Hier zeigt sich über einen längeren Zeitraum also eine deutliche Steigerung des Sprechtempos.

### *Dynamische Merkmale:*

Heath Ledger wird im Vergleich mit etwas geringerer Sprechlautstärke wahrgenommen als der deutsche Sprecher, dessen Lautheit als normal eingestuft wird. Außerdem variiert Heath Ledger auch etwas stärker in der Lautstärke, ein Merkmal, das oft mit einer sachlichen Sprechweise in Verbindung gesetzt wird (vgl. Zilliken 1991: 36). Dies würde zu Jokers Anliegen passen, ruhig und sachlich seine Motivation erklären zu wollen.

### *Melodische Merkmale:*

Die deutsche Stimme wird im Vergleich höher in der Tonhöhe eingeschätzt, auch erreicht dieser Sprecher Variationen innerhalb seiner Äußerungen mehr über die melodischen Merkmale. Trotzdem wird auch die englische Stimme mit einem weiten Tonhöhenbereich wahrgenommen, denn beide Sprecher bleiben nach einem Lachenfall des Jokers kurzzeitig im Falsett.

### *Artikulatorische Merkmale:*

Segmentlängungen finden vor allem bei Vokalen statt, jedoch bestehen hier auch sonst keine auffälligen Unterschiede.

### *Stimmklang:*

Beide Stimmen werden als leicht knarrend und leicht behaucht (+1) wahrgenommen, bei der englischen ist zeitweise auch eine raue Anregung zu hören, die der deutsche Sprecher nicht erreicht. Klangfülle und Klangfarbe sind in dieser Szene auditiv schwer zu bestimmen, wohl weil beide Sprecher hier viel variieren. Generell gleicht sich der Synchronsprecher aber sehr an das Original an.

### *Sprechspannung:*

Beide Stimmen zeigen leichte Nasalität, der deutsche Sprecher hat generell mehr Sprechspannung während der englische eher als normal gespannt wahrgenommen wird.

### 3.3.6 Hospital-Szene

In dieser Szene besucht der Joker, dem es inzwischen gelungen ist, aus dem Polizeihauptquartier zu fliehen, den Staatsanwalt Harvey Dent im Krankenhaus. Dieser wurde bei seiner Entführung durch eine Explosion schwer verletzt, konnte aber von Batman gerettet werden. Leider kam für Dents Freundin Rachel Dawes, die der Joker ebenfalls entführen ließ, jede Hilfe zu spät. Der einst korrekte Anwalt Harvey Dent ist dadurch schwer demoralisiert und will sich am Joker rächen, da er ihn für seine Lage verantwortlich macht. Der Joker nutzt nun diese hilflose Lage Harvey Dents dazu, ihm wie vorher Batman ebenfalls zu erklären, dass er ihm lediglich vor Augen führen wollte, dass Kontrolle und Planung eine Illusion sind und in Wahrheit das Chaos die Welt regiert. Analysiert wurden hier die Sequenzen TC 1:43:27 bis 1:44:35 und – als szenische Ergänzung – TC 1:44:51 bis 1:44:58.

#### *Temporale Merkmale:*

Der deutsche Sprecher wird ebenfalls wieder als schneller sprechend wahrgenommen, obwohl nun die AR mit 3,5 (englisch) zu 3,8 Silben pro Sekunde (deutsch) im Schnitt kaum schneller ist. Wieder machen hier vor allem die Pausen den Unterschied aus: Der deutsche Sprecher hat mit 5,2 Sekunden Differenz insgesamt deutlich weniger Pausen, im Vergleich bedeutet dies mit 31,5 % sogar den zweithöchsten Unterschied an Pausenzeit.

#### *Dynamische Merkmale:*

Heath Ledgers Sprechweise wird hier als im Vergleich leiser mit weniger Variation wahrgenommen, während der deutsche Sprecher hier ohne besondere Auffälligkeiten bleibt.

#### *Melodische Merkmale:*

Die englische Stimme wird in der Tonhöhe als niedriger eingeschätzt, mit weitem Tonhöhenbereich und wahrnehmbarer Tonhöhenvariation. Die deutsche Stimme ist dagegen in der Tonhöhe normal, weist allerdings ebenfalls einen recht weiten Tonhöhenbereich auf und zeigt darüber hinaus mehr Tonhöhenvariation.

#### *Artikulatorische Merkmale:*

Hier ebenfalls keine auffälligen Unterschiede.

#### *Stimmklang:*

Beide Stimmen zeigen eine knarrende Anregung, wobei die deutsche Stimme im Vergleich als sehr knarrend eingeschätzt wird. Außerdem gesellt sich im Englischen

zeitweise noch eine leicht raue Komponente hinzu. Beide Sprecher nutzen erneut die Falsettstimme, was die Wahrnehmung eines recht weiten Tonhöhenbereichs unterstützt.

#### *Sprechspannung:*

In der Nasalität zeigen beide Stimmen keinen Unterschied, dennoch weist der deutsche Sprecher vor allem supralaryngal und laryngal mehr Spannung auf, hier wird der englische eher als normal bis entspannt wahrgenommen. Zusammen mit der geringeren Lautstärke und weniger Variation könnte dies der entspannt-berichtenden Haltung des „englischen“ Jokers in dieser Szene entsprechen.

### 3.3.7 Zusammenfassung

Auffällige Unterschiede zwischen beiden Stimmen zeigen sich in allen analysierten Szenen in den temporalen Merkmalen. Die deutsche Stimme weist in der auditiven Wahrnehmung insgesamt eine höhere Sprechgeschwindigkeit auf. Dies ist vermutlich vor allem der Tatsache geschuldet, dass der deutsche Sprecher deutlich weniger Pausen macht, um auch die Synchronität der Szenen in Bezug auf die Gesamtdauer zu gewährleisten. Dadurch folgen die einzelnen Phrasen schneller aufeinander, was beim Hörer zur Wahrnehmung einer höheren Sprechgeschwindigkeit und damit zu einem anderen Eindruck führt (vgl. Maier 1997: 96-97), obwohl die AR beider Sprecher durchweg nahezu gleich ist. Die zwischen den Sprechern in einer Szene konstante AR bei viel höherer Silbenanzahl mag vor allem daher rühren, dass der Synchronsprecher die geforderte quantitative Lippensynchronität einhalten muss, d.h. es darf nur so lange gesprochen werden, wie auch Artikulationsbewegungen zu sehen sind (vgl. Maier 1997: 96). Diese Anpassung wird auch daran deutlich, dass die AR des Synchronsprechers im Vergleich nicht dem Schnitt eines deutschen Sprechers entspricht (vgl. Bose 2003: 141). Aus dem selben Grund ergeben sich wohl auch für beide Sprecher keine quantitativen Unterschiede in der Pausensetzung.

Diese Beobachtungen waren in nahezu allen Szenen konstant, außer in der sehr kurzen Vater-Szene, die ein umgekehrtes Bild präsentiert. Generell lässt sich etwa in längeren Szenen die AR deshalb leichter bestimmen, weil diese einen Durchschnittswert darstellt, der bei längerer Sprechzeit umso einheitlicher ausfällt, da dann eventuelle Abweichungen einzelner Sequenzen nicht so stark ins Gewicht fallen. Je kürzer eine Sprechersequenz, desto weniger aussagekräftig ist somit ihre AR bezogen auf das Gesamtkorpus.



Besonders auffällig sind auch die Unterschiede in den melodischen Merkmalen. Die deutsche Stimme wurde fast durchweg als leicht höher, mit einem breiteren Tonhöhenbereich und mehr Tonhöhenvariationen eingestuft. Dies konnte durch die akustischen Messungen weitgehend bestätigt werden, wenn auch die Unterschiede der Messwerte hin und wieder nicht so deutlich waren, da diese nur einen kleinen Teil des Wahrnehmungskomplexes *Tonhöhe* abbilden. So liegt auch Heath Ledgers Stimme mit einer durchschnittlichen Grundfrequenz von 153,4 Hz in den analysierten Szenen eher im mittleren Bereich, wirkt aber im Vergleich dennoch tiefer, obwohl dies nur etwa 20 Hz weniger sind als beim deutschen Sprecher (durchschnittliche Grundfrequenz 173,6 Hz). Generell weist der deutsche Sprecher in allen Szenen trotz desselben Kerninhalts der Äußerungen viel mehr Melodiebewegungen und -sprünge auf. Dass diese Unterschiede nicht sprachlicher sondern sprecherischer Natur sind, wird daran deutlich, dass er den globalen Melodieverlauf – wie etwa eine Abwärtsbewegung – in der Regel angleicht. Dies ist auch in den Grundfrequenzverläufen als Vergleichswert sichtbar (vgl. etwa Abb. 3.5 und 3.6). Bezogen auf den  $F_0$ -Range fällt außerdem auf, dass Heath Ledger in den längeren Szenen mehr Konstanz aufweist als der deutsche Sprecher, der dagegen vor allem in der Gordon-Szene einen sehr großen Range verwendet. Dies zeigt sich auch dadurch, dass er im Durchschnitt die kleineren  $F_0$ -Minima erreicht. Lediglich in der Video-Szene und der sehr kurzen Vater-Szene gleicht er sich im Range der Vorlage an. Hier wäre die Aussagekraft allerdings besser, wenn auch akustisch mehr längere Szenen hätten analysiert werden können, um zu sehen, ob dort Heath Ledgers Wert für den Range tatsächlich konstant bei etwa 170 liegt.

Hinsichtlich der dynamischen Merkmale ergaben sich keine klaren Unterschiede zwischen beiden Stimmen. Gerade wegen der unterschiedlichen Aufnahmebedingungen (vgl. Kapitel 3.2.3) ist hier aber auch durch akustische Messungen keine eindeutige Aussage möglich (vgl. auch Paeschke/Sendlmeier 1997: Abschn. 3). Lediglich die dynamische Akzentuierung fällt bei Heath Ledger gelegentlich etwas stärker aus, was aber vor allem vom szenischen Kontext abhängt.

Artikulatorisch waren die Unterschiede zu gering, um mögliche Abweichungen in der Eindrucksbildung zu erklären. Ein Grund hierfür ist sicherlich die in der Synchronisation geforderte Lippensynchronität. Auffällige Abweichungen zwischen Originalbild und deutschem Ton, wie etwa normales Sprechen bei gleichzeitig sichtbarer Nahrungsaufnahme, würden die Einheit von Bild und Ton empfindlich

stören. Deshalb muss der deutsche Sprecher auch die Artikulationsmodi hörbar an das Original angleichen (vgl. hierzu auch Maier 1997: 97f., Eckert/Laver 1994: 143).

Im Stimmklang zeigten beide Sprecher innerhalb einer Szene häufig zu viele Variationen, als dass sich eindeutige Unterschiede herausarbeiten ließen. Außerdem waren sich beide Sprecher gerade bezüglich ihrer laryngalen Settings sehr ähnlich, wenn auch deren wahrnehmbare Ausprägung hin und wieder leicht variierte. Insgesamt sind aber auch hier keine klaren Tendenzen erkennbar. Lediglich in Bezug auf die Klangfarbe wurde die Synchronstimme meist heller eingeschätzt, was bekanntermaßen aber auch mit einer höheren Stimmlage korreliert und nicht klar davon getrennt werden kann (vgl. Kranich 2002: 19, Bose 2003: 41).

Ähnlich sieht das Bild in Bezug auf die Sprechspannung aus: Zwar fällt auf, dass beide Sprecher in den einzelnen Spannungen lokal unterschiedlich variieren (so setzt Heath Ledger etwa in der Gordon-Szene mehr Variation in der Sprechspannung als in der Melodik ein), was auch erwartungsgemäß sehr gut wahrgenommen wird (vgl. Paeschke/Sendlmeier 1997: Abschn. 3.1.1), aber gerade diese Variationen erschweren eine eindeutige auditive Beurteilung. Deutliche globale Unterschiede treten nicht auf.

Somit ist ersichtlich, dass beide Stimmen sich in den analysierten Szenen hauptsächlich bezüglich des temporalen Verlaufs, der melodischen Merkmale wie Tonhöhen-/ $F_0$ -Variation und -Range sowie Klangfarbe mehr oder weniger stark unterscheiden. Andere Merkmale wie Sprechspannung und Stimmklangvariation wirken wohl als Komplex am Eindruck mit, zeigen jedoch zu wenig deutliche und eindeutig klassifizierbare Unterschiede.

Dies erhärtet aber dennoch die Vermutung, dass beide Stimmen in einzelnen Merkmalsausprägungen und damit wohl in ihrem Gesamtausdruck nicht identisch sind. Die Voraussetzung für die Hypothese dieser Arbeit konnte somit bestätigt werden. Inwieweit diese Unterschiede jedoch möglicherweise die unterschiedliche Wirkung auf die Internetnutzer erklären können oder ob womöglich weitere Ursachen mit hineinspielen, soll im folgenden Kapitel diskutiert werden.

## 4. Diskussion

### 4.1 Ergebnisdiskussion im Hinblick auf bisherige Ergebnisse

Aus den Hörerurteilen, die in Kapitel 3.2.2 dargestellt wurden, geht hervor, dass der deutschen Stimme das Merkmal *verrückt/wahnsinnig/irre/psycho* deutlich mehr zugeschrieben wird, als der Originalstimme, nämlich in der Hälfte aller abgegebenen Eindrücke. Gleichzeitig wird die Originalstimme eher als böse, gruselig und sogar allein als tief bzw. dunkel und rauh wahrgenommen. Somit liegt der Schluss nahe, dass diese Eindrücke nicht nur willkürlich sind, sondern durch die unterschiedliche Ausprägung stimmlicher Merkmale zustande kommen. Für die im Vergleich tiefere Lage von Heath Ledgers Stimme sprechen vor allem die auditiven, aber auch die akustischen Analyseergebnisse. Außerdem hat die Analyse im vorangegangenen Kapitel ergeben, dass sich beide Stimmen vor allem in den temporalen und melodischen Merkmalen unterscheiden. Dies konkretisiert die Hypothese dieser Arbeit folgendermaßen: Wenn die Unterschiede in der Sprechwirkung tatsächlich auf diese stimmlichen Unterschiede zurückzuführen sind, müssten sie sich auch durch diese erklären lassen. Um diese Annahme abschließend zu verifizieren, wird nun im Folgenden versucht, die Ergebnisse der eigenen Analyse durch bisherige Erkenntnisse der Sprechwirkungsforschung mit den unterschiedlichen Eindrucksausprägungen in Beziehung zu setzen.

Schon Fährmann war es in seiner Analyse gelungen, bestimmte stimmliche Merkmalsausprägungen mit einzelnen wahrgenommenen Persönlichkeitseigenschaften in Verbindung zu bringen. Dazu fasst er im Anschluss an seine Untersuchung jeweils bestimmte stimmliche Manifestationen zu einzelnen seelischen Zustandsgruppen zusammen.

Eine ausgeprägte Melodiebewegung gepaart mit erhöhtem Tempo, wie sie in dieser Arbeit bei der Synchronstimme zu beobachten war, steht dabei laut Fährmann einmal für einen gewissen aktiven Willen (vgl. Fährmann 1960: 155), sowie allgemein einen höheren Grad an Erregung seitens des Sprechers (vgl. Fährmann 1960: 61). Andererseits werden beide in extremer Ausprägung – insbesondere bei hoher Tonhöhenänderungsfrequenz – als „allgemeine Zeichen des nervösen Syndroms“ (Fährmann 1960: 165) angesehen. Diese Erkenntnis ist für Fährmann relativ gesichert, da die menschliche Stimme erfahrungsgemäß sehr sensibel auf seelische Regungen

reagiere und diese transportiere. Das heißt, hiernach lassen weite, häufig wechselnde melodische Bewegungen gepaart mit sehr raschem, gehetztem Tempo oft auf eine gewisse seelisch instabile, wenn nicht „wahnsinnige“ Komponente in der Sprecherpersönlichkeit schließen. Dies legte in der vorliegenden Analyse den Schluss nahe, dass wohl das erhöhte Tempo und die ausgeprägteren Melodiebewegungen des deutschen Sprechers, die auch am wechselnden Range sichtbar werden, das Merkmal Wahnsinn/Psycho als Ausdruck einer nervösen Störung mehr transportierten, als Heath Ledgers eigener Sprechausdruck, der nur gelegentlich starke melodische Sprünge einsetzt (vgl. etwa die Beispiele in Kapitel 3.3.1 und 3.3.4). Fraglich wäre allerdings, wo hier das Element der Komik mit hineinspielt, welches drei Hörer ebenfalls wahrzunehmen meinen.

In Bezug auf die Merkmale *böse* oder *gruselig/furchteinflößend*, mit welchen fast ausschließlich die Originalstimme in Verbindung gebracht wird, formuliert Fährmann in der Gruppe „Hass, Grausamkeit, Bösartigkeit“ folgende stimmlichen Eigenschaften: eine „harte“, voluminöse Stimme, hohe Sprechspannung auch in den Tonhöhenbewegungen, gepaart mit einer schärferen Artikulation (Fährmann 1960: 167). Dies würde sich mit der auditiven Beobachtung decken, dass Heath Ledger vor allem in den längeren Szenen stark lokal mit der Sprechspannung und im Stimmklang variiert, während der deutsche Sprecher hier wiederum Variation über die Melodiebewegungen erreicht. Das heißt, der Merkmalskomplex aus mehr Variation in der Sprechspannung mit gleichzeitig weniger melodischer Variation bei dunklerer und tieferer Stimmlage könnte im direkten Vergleich beider Stimmen bewirken, dass Heath Ledgers Stimme aggressiver wirkt. Auch Rauheit war ja ein Merkmal erhöhter Sprechspannung.

Allerdings sind Fährmanns scheinbar so klare Zuordnungen bei genauerer Betrachtung mit Vorsicht zu genießen. Denn zum einen benutzt er sehr häufig schwer zuordenbare Formulierungen für bestimmte Stimmklangmerkmale, die wenig mehr aussagen als eine subjektive Eindrucksbeschreibung und nicht klar an objektiv nachvollziehbaren Parametern festzumachen sind, wie dass eine Stimme „ölig, weich, schlapp“ (Fährmann 1960: 162) klinge. So werden unscharfe Adjektive aus anderen Erfahrungsbereichen mit anatomischen Stimmqualitätsmerkmalen vermischt, was einen Vergleich mit anderen Untersuchungsergebnissen erschwert (vgl. auch Kranich 2002: 19). Des Weiteren ist Fährmann (1960: 155) davon überzeugt, dass „[w]enn wir einzelne, für den Sprecher bedeutsame seelische Wesenszüge klar erfasst [...] haben, dann ist es nicht schwer, die richtigen Sprechmerkmale zuzuordnen“. Diese Herangehensweise ist jedoch eher

problematisch. Davon abgesehen, dass es selten „richtige“, sondern immer nur „mehr oder weniger wahrscheinliche“ Sprechmerkmale eines bestimmten Gefühlszustandes o.ä. gibt (vgl. auch Scherer 1982: 201), erweckt diese Aussage, mit Erinnerung an die Ausführungen von U. Geißner (vgl. Kapitel 2.2.1), doch den Verdacht, dass hier, ausgehend von bestimmten, aufgrund von Wahrnehmungen vorgenommenen, Interpretationen der Persönlichkeit, einzelne Merkmale möglicherweise ebenso in den Sprechausdruck „hineininterpretiert“ wurden. Denn man wusste ja aufgrund der seelischen Einschätzung der Person schon, was man anschließend in den einzelnen Merkmalen des Schallereignisses zu hören hatte. Damit wäre man hier wieder bei der Tatsache angelangt, dass „the judges tended to stereotype the personality of the speaker according to the sound of his voice“ (Addington 1968: 10), nur mit umgekehrtem Vorzeichen. Diese subjektive Stereotypisierung würde ebenfalls keine auf andere Situationen übertragbaren Aussagen erlauben. Das heißt, von Fährmanns Ergebnissen lässt sich möglicherweise zu denen der eigenen Analyse nur schwer ein gültiger Zusammenhang herstellen. Man benötigt hier somit weitere Daten aus anderen Untersuchungen, um die oben vermuteten Zusammenhänge wahrscheinlicher zu machen.

Obwohl Fährmanns Ergebnisse in ihrer Aussagekraft fragwürdig sind, stellen Stock und Suttner (1990: 61) in eigenen Untersuchungen ebenfalls fest, dass  $F_0$ -Range und – Kontur zusammen mit einer erhöhten Sprechgeschwindigkeit signifikante Effekte bei der Beurteilung von Stimme als erregter aufweisen.

Auch Zilliken (1991: 30) weist in ihrer Zusammenfassung verschiedener Untersuchungen auf eine positive Korrelation zwischen Stimmumfang und einer wahrgenommenen Erregung hin. Außerdem wirke ein erhöhtes Sprechtempo ebenfalls affektvoller. Dieser Eindruck von Erregung werde noch verstärkt, wenn im deutschen Sprechpausen zu kurz sind oder nicht am logischen Sinnschritt eingesetzt würden, ihre klare, gliedernde Funktion somit verloren ginge (vgl. Zilliken 1991: 33f.). Die im Vergleich zwischen beiden Stimmen im Deutschen erheblich kürzeren Pausen könnten genau diese Einschätzung hervorgerufen haben, weil die einzelnen Phrasen für den Hörer unwillkürlicher aufeinander folgten als im Original. Außerdem entsteht durch größere Tonhöhendifferenzen ebenfalls „der perzeptive Eindruck einer größeren Sprechgeschwindigkeit“ (Redecker 2008: 152).

Daneben bringt die Autorin ein unregelmäßiges Steigen und Fallen der Melodie ebenfalls mit dem Ausdruck einer labilen Persönlichkeit in Verbindung (vgl. Zilliken 1991: 38). In Kombination könnten diese beiden Merkmale also auch hier einen nervös-verrückten Eindruck verstärken.

Scherer (1982: 198; vgl. auch Kapiel 2.3) folgert aus den eigenen Analyseergebnissen ebenfalls, „daß die Grundfrequenz der Stimme möglicherweise ein valider Indikator für affektive Störungen ist“. So wird eine höhere Grundfrequenz ebenfalls mit Erregung oder Stress in Verbindung gebracht, was besonders im Vergleich zum Normalzustand eine recht zuverlässige Einschätzung zu sein scheint. Ein Nachteil in der Vergleichbarkeit von Scherers Ergebnissen ist sicherlich, dass hier inhaltsfreies Stimulusmaterial verwendet wurde, wodurch eine stärkere Fokussierung auf stimmliche Merkmale stattfand, als dies bei „sinnvollen“ Äußerungen möglicherweise der Fall ist (vgl. Scherer 1982: 197 u. 199).

Zusätzlich betonen Eckert und Laver in ihrer Zusammenfassung verschiedener Ergebnisse folgende Zusammenhänge: Generell wird jede stimmliche Normabweichung von Hörern mit eher negativen Eigenschaften assoziiert (vgl. Eckert/Laver 1994: 161). So könnte tatsächlich auch die manchmal höhere Stimmlage und die abweichende Melodisierung des deutschen Sprechers gepaart mit einem zu schnellen Tempo eher auch zu Attributionen des Abnormalen wie Wahnsinn oder Verrücktheit führen, als dies bei der Originalstimme der Fall ist. Denn bei dieser treten möglicherweise dafür Merkmale des aggressiven Sprechens wie rauher Stimmklang und hohe Sprechspannung in den Vordergrund (vgl. Eckert/Laver 1994: 163f.) und beeinflussen so den stimmlichen Komplexeindruck stärker als beim deutschen Sprecher.

Daneben konnte Redecker (2008: 147) anhand einer vergleichenden Analyse der Wirkung eines Werbespots erneut zeigen, „dass bereits geringe stimmliche und sprecherische Modifikationen innerhalb weniger Sekunden eine hochsignifikante Veränderung in der Wahrnehmung und in der Wirkung des Werbespots nach sich ziehen“, auch bezüglich der darin dargestellten Person. Somit wäre durch diese Daten tatsächlich die Hypothese dieser Arbeit bestätigt.

Dennoch zeigt eine genauere Betrachtung, dass der Bezug zwischen den Nutzerurteilen und den in der vorliegenden Analyse festgestellten Unterschieden zwischen den Stimmen nicht so einfach hergestellt werden kann.

Zilliken (1991: 36) etwa merkt an, dass es hier gelte, „Vorsicht walten zu lassen, da menschliche sprachliche Ausdrucksformen sich als zu komplex und vielfältig darbieten, als daß sich einzelne Sprechausdrucksmerkmale in entsprechende „Schubladen“ einordnen ließen“. Auch der genauere Vergleich der von Hörern zugeschriebenen Eigenschaften mit denen stabiler Bekannteneinschätzungen in Scherers Daten zeigt, dass situativ sehr valide stimmliche Hinweisreize auf keinen Fall auf stabile oder habituelle Merkmale übertragen werden können (vgl. Scherer 1982: 202). Dies lässt erneut darauf schließen, dass auch sehr konstante überindividuelle Einschätzungen situationsabhängig sind und nicht ohne weiteres auf andere übertragen werden können, wenn sich auch gewisse Tendenzen stets gleichen. Denn auch wenn erhöhte Melodiebewegung und höheres Tempo somit von einigen Autoren mit Erregtheit und nervösen Störungen in Verbindung gebracht werden (vgl. auch Kehrein 2002: 323, Redecker 2008: 152) und im Gegensatz dazu eine gespannte Stimme mit wenig Melodik Aggressivität oder Drohung bedeutet, gilt es dennoch nochmals zu beachten, dass Wahrnehmung und Interpretation nie kontextfrei geschieht. Zum Gesamtbild zählen alle im Augenblick der Eindrucksbildung vorhandenen Teilbedeutungen der verschiedenen Signalisierungssysteme, es gibt nicht *die* Ausdrucksbedeutung (vgl. etwa Kehrein 2002: 321). Um deshalb sichere Aussagen über Stimmwirkung treffen zu können, müsste die Stimme auch alleiniges Unterscheidungskriterium sein. Dies kann durch die vorliegende Untersuchung jedoch nur sehr eingeschränkt gewährleistet werden, da sie methodisch neben einem Vorteil auch deutliche Nachteile aufweist.

Der Vorteil ist sicherlich, dass die Beurteiler zum Zeitpunkt ihrer Urteilsabgabe nicht wussten, dass diese Gegenstand einer Untersuchung sein würden. Deshalb ist zu erwarten, dass die spontan formulierten stimmlichen Eigenschaften nicht durch Versuchs- und Versuchsleitereffekte oder übermäßige Reflexion verzerrt sind (vgl. Kapitel 1 u. Kapitel 3.2.2). Dies war allein dadurch möglich, weil als Grundlage für die Urteile ein öffentlich zugängliches Kommentarforum des Internets diente, dessen vorrangiger Zweck nicht die Beurteilung von Stimmen, sondern die Bewertung und Kommentierung einzelner Videos ist. Um so erstaunlicher ist die Tatsache, dass zwischen manchen Eigenschaften – insbesondere in Bezug auf die Synchronstimme – ein recht breiter Konsens besteht (vgl. Kapitel 3.2.2).

Ebenso darf die Beurteilergruppe aber in Bezug auf Alter, Geschlecht, sozialer Hintergrund, Bildungsstand etc. nicht zu heterogen sein (vgl. Eckert/Laver 1994: 151ff.,

Stock 1991: 58). Daneben spielen außerdem äußere Bedingungen wie Tageszeit und Ort beim Beurteilungsverhalten eine Rolle (vgl. Stock 1991: 57).

Hierin liegt auch der große Nachteil der hier verwendeten Eindruckserhebung. Denn dadurch, dass auf weitgehend anonyme Kommentare aus dem Internet zugegriffen wurde, war es nicht möglich, die Beurteilergruppe bezüglich dieser Faktoren konstant zu halten. Deshalb ist es notwendig, die Nutzerurteile bezüglich folgender Punkte zu relativieren:

1. Es ist weder klar, in welcher Situation die einzelnen Nutzer das Video oder die DVD sahen, noch ob die Darbietung jeweils bezüglich der Tonqualität konstant war. Dies spielt bei der Eindrucksgewinnung eine nicht unerhebliche Rolle, da je nach Darbietungsart z.B. Störschall die Beurteilung beeinflussen kann (vgl. Bose 2003). Was hier hingegen nicht so stark ins Gewicht zu fallen scheint, ist die Tatsache, dass nahezu der gesamte Film mit Musik unterlegt ist. Denn Redecker stellte bei ihrer Analyse etwa fest, „dass zusätzliche Stimuli wie Musik und Bild die Wirkung prosodischer Impulse nicht nivelliert haben [...]“ (Redecker 2008: 151).
2. Eine weitere Einschränkung der Eindeutigkeit der Untersuchungsergebnisse ergibt sich aus der Tatsache, dass die Nutzer mit dem Titel des Videos schon explizit zu einem stimmlichen Vergleich aufgefordert werden, was eine klare Selektion der Wahrnehmung darstellt (vgl. U. Geißner 1984: 129). Außerdem ist nicht klar, wie oft die Nutzer sich die dargebotene Szene „angehört“ oder den Film gesehen haben, bevor sie ihren Kommentar abgaben. Dies kann möglicherweise trotzdem zu verzerrenden Eindrücken führen, die unter normalen Umständen einer Filmpräsentation nicht entstanden wären.
3. Des Weiteren sind die Eindrucksbeschreibungen zwar spontan, dadurch aber auch schwer nachvollziehbar und vergleichbar. Denn es ist zum Beispiel nicht klar, ob unterschiedliche Adjektive wie „wahnsinnig“ und „psycho“ tatsächlich dasselbe Konzept meinen, auch wenn sie es im allgemeinen Verständnis vielleicht beinhalten. Dasselbe gilt für die schwer zu fassenden Beschreibungen des Stimmklangs. Denn aus der vorliegenden Analyse geht nicht klar hervor, warum Heath Ledgers Stimme als rauer beurteilt wird. Ein weiterer Nachteil der spontanen Äußerungen ist, dass sich die Nutzer häufig nur zu einer Stimme



äußern und die so entstandenen Aussagen sehr viel Interpretationsspielraum in Bezug auf den Stimmenvergleich zulassen.

4. Abschließend ist zu bedenken, dass auch stimmliche Eindrücke in der Regel in direktem Bezug auf die Handlungen der betreffenden Person gebildet werden (vgl. U. Geißner 1984: 134). Das heißt, Eigenschaften, die man jemandem aufgrund seiner Handlungen zuschreibt, wird man auch in dessen stimmlichem Ausdruck zu bestätigen suchen. Das bedeutet in dem vorliegenden Fall, dass dadurch, dass die Nutzer die Handlungen und Verhaltensweisen bereits aus den Comics und Vorgängerfilmen kennen, sie zu wissen meinen, wie dessen Stimme zu klingen hat, und möglicherweise eher geneigt sind, schon kleinste stimmliche Unterschiede dahingehend zu interpretieren. Somit ist nicht klar, ob dieser Ausdruck für sich genommen ebenfalls so interpretiert worden wäre.
5. Auch kulturelle Unterschiede können zu diesem Effekt führen. Denn wie schon in Kapitel 2.2.1 ausgeführt, können Sprechausdrucksmuster als solche nur dann richtig gedeutet werden, wenn sie innerhalb einer Sprachgemeinschaft konventionalisiert sind, d.h. wenn auch entsprechende Hörmuster existieren (vgl. Geißner 1984a: 25). Kommt es zur Überschneidung mit anderen Sprachgemeinschaften, so werden die eigenen Muster auf die der fremden Sprache übertragen. Deshalb wird eine normale amerikanische Männerstimme in Deutschland in der Regel als „zu tief“ eingestuft, obwohl sie sich für Amerikaner noch innerhalb der Norm befindet (vgl. Eckert/Laver 1994: 156, Redecker 2008: 34 u. 147). Dies könnte auch bei der Beurteilung der Stimme Heath Ledgers durch die überwiegend deutschen Nutzer mit eine Rolle gespielt haben, weshalb der Eindruck der im Vergleich „tiefen“ Stimme Ledgers nicht allein aus einem Tonhöhenvergleich abgeleitet werden kann. Bezogen auf die Grundfrequenz liegen beide Stimmen ja eher im mittleren Bereich, auch wenn Ledgers Stimme generell ein geringeres  $F_0$  aufweist.

Diese Einschränkungen machen die Nutzerurteile zu nicht reinen Eindrücken und sind bei einer abschließenden Interpretation der Analyseergebnisse zu beachten.

## 4.2 Schlussfolgerungen und Ausblick

Ziel der Arbeit war es, durch eine akustische und auditive Analyse unterschiedliche spontane Eindrücke von Internetnutzern zwischen der Originalstimme von Heath

Ledger und seiner Synchronstimme nachvollziehbar zu machen bzw. zu überprüfen, ob diese Eindrücke tatsächlich eine stimmliche Grundlage haben.

Die Analyse konnte zeigen, dass vor allem in Bezug auf den Komplex der melodischen und temporalen Parameter recht deutliche wahrnehmbare und messbare Unterschiede zwischen beiden Stimmen bestehen, trotz der Tatsache, dass nur wenige Teile des Films tatsächlich als Grundlage genutzt werden konnten. Für die anderen Parameter konnten zum Teil keine klaren allgemeinen szenenübergreifenden Aussagen getroffen werden. Dennoch gilt es zu beachten, dass diese stimmlichen Unterschiede nicht per se die spontanen Eindrücke der Internetnutzer erklären können, da Wahrnehmung und Sprechwirkung immer als Komplex aus stimmlichen, situativen, individuellen o.ä. Merkmalen geschieht. Das heißt, es wäre falsch, die Ergebnisse dahingehend zu interpretieren, dass etwa eine stärkere Variation der Grundfrequenz, sowie ein weiterer Range und damit auch häufige wahrnehmbare Änderungen in der Tonhöhe gepaart mit schnellerem Tempo ganz klar den Eindruck „verrückt“ alleine hervorrufen. Sondern dies bedeutet, dass es zumindest nicht unwahrscheinlich ist, dass im Vergleich beider Stimmen die Ausprägtheit dieses Unterschieds eine nicht unwesentliche Rolle bei der Interpretation des stimmlichen Eindrucks durch die Nutzer in Richtung „verrückt“ spielt. Denn diese stimmlichen Merkmale scheinen eher dem Rollenbild des „Psycho-Clowns“ zu entsprechen, das die meisten Nutzer von vornherein von der Figur des Jokers haben. Immerhin deuten auch andere Ergebnisse in diesem Zusammenhang auf einen dominanten Einfluss der melodischen und temporalen Parameter hin. Es scheint also, dass die vorliegende Analyse zumindest einen Teil der unterschiedlichen Eindrücke tendenziell nachvollziehbar machen kann, da gezeigt werden konnte, dass sie zu einem Teil auch auf stimmliche Unterschiede zurückgehen können. Dies ist umso wahrscheinlicher, da auch für die Internetnutzer der stimmliche Unterschied die jeweils dominante Variable war.

Außerdem zeigt diese Analyse damit, dass schon minimale Änderungen in den Teilen des Gesamtkomplexes Stimme einen unterschiedlichen Eindruck einer Synchronisation hervorrufen können und die Zuschauer dies wahrnehmen. Damit trägt auch die Synchronisation ihren Teil zur Vermittlung des Rollenbildes bei und kann die Gesamtwirkung eines Films mit beeinflussen. Denn etwa durch eine nicht kongruente Besetzung der Stimme können manche Wirkungsaspekte des Originals verloren gehen, wodurch das ausländische Publikum das Werk – wenn auch nur minimal – anders aufnimmt als das des Ursprungslandes. Ob dies von den Verantwortlichen immer

gewollt ist, sei dahingestellt. So hat sich gezeigt, dass die durch ein erhöhtes Sprechtempo hervorgerufenen Eindrücke in der Regel nicht intendiert, sondern allein der technischen Notwendigkeit der Synchronität geschuldet sind (vgl. auch Maier 1997: 96f.). Gerade weil aber im Zeitalter der DVD immer mehr Menschen den direkten Stimmenvergleich vornehmen können, sollte sich die Synchronbranche mehr dessen bewusst sein, um die Illusion der Einheit von Bild und Ton aufrechterhalten zu können. Methodisch interessant sind die Ergebnisse dieser Arbeit, weil sie zeigen, dass Hörer auch in einer Nichtlaborsituation durchaus in der Lage sind, bis zu einem gewissen Grad überindividuell nachvollziehbare stimmliche Eindrücke zu erlangen und dass diese nicht völlig inkonsistent und willkürlich sind.

Trotzdem bleibt das Problem, dass nicht klar gesagt werden kann, ob diese Eindrücke auch von anderen Gruppen so vorgenommen worden wären, da es sich um eine recht kleine, speziell an der Thematik interessierte Gruppe von Beurteilern handelt, die auch keine klaren, unabhängigen Daten zu jeder der beiden Stimmen liefern. Hier bestünde die Möglichkeit, dies durch eine umfassendere Studie unter Laborbedingungen mit einer größeren Hörergruppe zu überprüfen. Käme diese zu ähnlichen Ergebnissen, so wäre es noch wahrscheinlicher, dass diese Eindrücke trotz der im vorangegangenen Kapitel genannten Einschränkungen auch eine stimmliche Grundlage haben.

Des Weiteren kann die vorliegende Analyse ebenfalls nicht zweifelsfrei klären, welche einzelnen stimmlichen Parameter nun in welchen Ausprägungen für die jeweiligen stimmlichen Eindrücke verantwortlich sind. Eine Möglichkeit, dies zu überprüfen, wäre, eine neutrale Äußerung beider Sprecher synthetisch in Bezug auf Merkmale wie Range und Sprechgeschwindigkeit zu verändern und dann jeweils mit entstandenen Eindrücken zu vergleichen. Ebenso scheint es sinnvoll, bei beiden Stimmen zusätzlich alle Parameter genauer in ihrer lokalen Veränderung zu untersuchen, da die vorliegende Analyse vermuten lässt, dass bei Merkmalen wie Sprechspannung und Stimmklang vor allem die Variation innerhalb einer Äußerung zum Gesamteindruck beiträgt und nicht so sehr eine global wahrnehmbare Ausprägung.

Insgesamt erscheint es mir dennoch angebracht, den Weg der spontanen Eindrücke weiterzuverfolgen, anstatt Beurteilergruppen feste Fragebögen ausfüllen zu lassen, die ihnen die zu hörenden Merkmale bis zu einem gewissen Grad schon vorgeben. Auch wenn deutlich wurde, dass wegen der verschiedenen methodischen Einschränkungen

offene Korpora aus dem Internet dafür nur eine mangelhafte Möglichkeit bieten, stellt die vorliegende Arbeit einen ersten Schritt in diese Richtung dar.

## 5. Literaturverzeichnis

- Addington, David W. (1968): *Voice and the Perception of Personality. An Experimental Study.* Oklahoma: Oklahoma State University Monographs.
- Banhold, Lars (2008, <sup>3</sup>2009): *Batman. Konstruktion eines Helden.* Bochum: Bachmann Verlag.
- Blaseio, Gereon (2003): „Gendered Voices“ in der Filmsynchronisation. „First Blood“ versus „Rambo“. In: Epping-Jäger, Cornelia/ Erika Linz (Hgg.): *Medien, Stimmen.* Köln: DuMont, S. 161-175.
- Bose, Ines (2003): *dóch da sín ja' nur mûster //.* Kindlicher Sprechausdruck im sozialen Rollenspiel (Hallesche Schriften zur Sprechwissenschaft und Phonetik 9). Frankfurt/Main u.a.: Peter Lang Verlag.
- Braun, Angelika/ Heilmann, Christa M. (2006): *Fremde Stimmen und fremde Körper. Zum Ausdruck von Emotionen im synchronisierten Film.* In: Bogner, Andrea et al. (Hgg.): *Jahrbuch Deutsch als Fremdsprache 31 (2005).* München: iudicum verlag, S. 164-189.
- Bräutigam, Thomas (2001): *Lexikon der Film- und Fernsehsynchronisation.* Berlin: Lexikon Imprint Verlag (Schwarzkopf & Schwarzkopf).
- Brown, Bruce L. (1982): *Experimentelle Untersuchungen zur Personenwahrnehmung aufgrund vokaler Hinweisreize.* In: Scherer, Klaus R. (Hg.), S. 211-227.
- Dath, Dietmar (2005): *Batman oder Ich bin der Ausnahmezustand.* In: Klage, Jan P. et al. (Hgg.): *Bob Kane Batman (Klassiker der Comic-Literatur 7).* Frankfurt/Main: Frankfurter Allgemeine Zeitung/Panini, S. 3-10.
- Eckert, Hartwig/ Laver, John (1994): *Menschen und ihre Stimmen. Aspekte der vokalen Kommunikation.* Weinheim: Psychologie VerlagsUnion (Beltz).
- Eckert, Hartwig (2004): *Atmung und Stimme.* In: Pabst-Weinschenk, Marita (Hg.): *Grundlagen der Sprechwissenschaft und Sprecherziehung.* München u.a.: Ernst Reinhardt Verlag, S. 20-31.
- Fährmann, Rudolf (1960): *Die Deutung des Sprechausdrucks. Studien zur Einführung in die charakterologische Stimm- und Sprechanalyse.* Bonn: Bouvier Verlag.
- Fischer, Peter-Michael/ Hartlieb, Karl (1984): *Die Stimme des Menschen. Aufbau, Funktion und Leistung.* Frankfurt/Main u.a.: Peter Lang Verlag.
- Franz, Katja (2008): *Kommunikative Aneignung von Fernsehserien und mediale Kommunikationskultur: AllyDE@yahoogroups.de (Inaugural-Dissertation).* Norderstedt: Books on Demand. Onlinequelle.  
URL: <http://archiv.ub.uni-marburg.de/diss/z2008/0664/pdf/dkf.pdf> (05.08.09).
- Geißner, Hellmut (1984a): *Über Hörmuster. Gerold Ungeheuer zum Gedenken.* In:

- Gutenberg, Norbert (Hg.): Hören und Beurteilen: Gegenstand und Methode in Sprechwiss., Sprecherziehung, Phonetik, Linguistik u. Literaturwiss. (Sprache und Sprechen 12). Frankfurt/Main: Scriptor, S. 13-56.
- Geißner, Hellmut (1984b): Funktionen des Sprechausdrucks in der Sinnkonstitution. In: Berger, Lothar (Hg.): Sprechausdruck (Sprache und Sprechen 13). Frankfurt/Main: Scriptor, S. 9-26.
- Geißner, Hellmut (1981, <sup>2</sup>1988): Sprechwissenschaft: Theorie der mündlichen Kommunikation. Frankfurt/Main: Scriptor.
- Geißner, Hellmut (1989): <o, oh> [o:]. Phonetisches – Prosodisches – Poetisches. In: Slembek, Edith (Hg.): Von Lauten und Leuten. Festschrift für Peter Martens zum 70. Geburtstag (Sprache und Sprechen 21). Frankfurt/Main: Scriptor, S. 69-82.
- Geißner, Ursula (1984): Hören und Beurteilen. ‚Wer Ohren hat zu hören, der höre.‘ (Mt. 13,43). In: Gutenberg, Norbert (Hg.), S. 127-147.
- Heilmann, Christa M. (2002): Interventionen im Gespräch. Neue Ansätze der Sprechwissenschaft. Tübingen: Niemeyer Verlag.
- Heilmann, Christa M. (2004): Was hört mein Ohr – was misst mein Computer?. In: Geissner, Hellmut K. (Hg.): Das Phänomen Stimme in Kunst, Wissenschaft, Wirtschaft (4. Stuttgarter Stimmtage 2002). St. Ingbert: Röhrig Universitätsverlag, S. 105-109.
- Hesse-Quack, Otto (1969): Der Übertragungsprozess bei der Synchronisation von Filmen. Eine interkulturelle Untersuchung. München: Ernst Reinhardt Verlag.
- Kehrein, Roland (2002): Prosodie und Emotionen. Tübingen: Niemeyer Verlag.
- Kienast, Miriam (2002): Phonetische Veränderungen in emotionaler Sprechweise. Aachen: Shaker.
- Kowal, Sabine (1991): Über die zeitliche Organisation des Sprechens in der Öffentlichkeit. Pausen, Sprechtempo und Verzögerungen in Interviews und Reden von Politikern. Bern u.a.: Verlag Hans Huber.
- Kranich, Wieland (2002): Suprasegmentale Eigenschaften gesprochener Sprache unter besonderer Berücksichtigung emotionaler Ausdrucksqualitäten. Jena: Univ.-Diss. Onlinequelle.  
URL: <http://www.db-thueringen.de/servlets/DerivateServlet/Derivate-1286/Dissertation.pdf> (29.05.09).
- Laver, John (1980): The Phonetic Description of Voice Quality. Cambridge u.a.: Cambridge University Press.
- Maier, Wolfgang (1997): Spielfilmsynchronisation (Forum Anglicum 23). Frankfurt/Main u.a.: Peter Lang Verlag.

- Niebuhr, Oliver (2007): *Perzeption und kognitive Verarbeitung der Sprechmelodie. Theoretische Grundlagen und empirische Untersuchungen (Language, Context, and Cognition 7)*. Berlin u.a.: Walter de Gruyter.
- Paeschke, Astrid/ Sendlmeier, Walter F. (1997): Die Reden von Rudolf Scharping und Oskar Lafontaine auf dem Parteitag der SPD im November 1995 in Mannheim. Ein sprechwissenschaftlicher und phonetischer Vergleich von Vortragsstilen. In: *ZfAL* 27/1997, S. 5-39. Onlinequelle.  
URL: [http://pascal.kgw.tu-berlin.de/home/publications/MA/ma\\_zfal.html](http://pascal.kgw.tu-berlin.de/home/publications/MA/ma_zfal.html) (05.08.09).
- Peitz, Christiane (2008): Wer zuletzt lacht (ersch. am 17.08.2008). *Der Tagesspiegel – Kultur – Comics*. Onlinequelle.  
URL: <http://www.tagesspiegel.de/kultur/comics/Batman-Christopher-Nolan;art18281,2594959> (05.08.09).
- Pétursson, Magnús/ Neppert, Joachim M. H. (1990, <sup>3</sup>2002): *Elementarbuch der Phonetik*. Hamburg: Buske Verlag.
- Pitzke, Marc (2008): Hollywood-Star Heath Ledger. Tod eines Herzensbrechers (ersch. am 23.01.2008). *SPIEGEL-ONLINE – Nachrichten – Kultur*. URL: <http://www.spiegel.de/kultur/kino/0,1518,530366,00.html> (05.08.09).
- Pruys, Guido Marc (1997): *Die Rhetorik der Filmsynchronisation. Wie ausländische Spielfilme in Deutschland zensiert, verändert und gesehen werden*. Tübingen: Gunter Narr Verlag.
- Redecker, Beate (2008): *Persuasion und Prosodie. Eine empirische Untersuchung zur Perzeption prosodischer Stimuli in der Werbung (Hallesche Schriften zur Sprechwissenschaft und Phonetik 25)*. Frankfurt/Main u.a.: Peter Lang Verlag.
- Runkehl, Jens/ Schlobinsky, Peter/ Siever, Torsten (1998): *Sprache und Kommunikation im Internet. Überblick und Analysen*. Opladen u.a.: Westdeutscher Verlag.
- Scherer, Klaus R. (Hg. 1982): *Vokale Kommunikation. Nonverbale Aspekte des Sprachverhaltens*. Weinheim u.a.: Beltz.
- Scherer, Klaus R. (1982): Stimme und Persönlichkeit – Ausdruck und Eindruck. In: Scherer, Klaus R. (Hg.), S. 188-210.
- Scherer, Klaus R. et al. (1982): Die Attribution von Persönlichkeitsmerkmalen aufgrund auditorischer und visueller Hinweisreize. In: Scherer, Klaus R. (Hg.), S. 228-252.
- Schulz von Thun, Friedemann (1981): *Miteinander reden: 1. Störungen und Klärungen. Allgemeine Psychologie der Kommunikation (Sonderausgabe 2008)*. Reinbek: Rowohlt Taschenbuch Verlag.
- Stock, Eberhard (1991): Grundfragen der Sprechwirkungsforschung. In: Krech,

Eva-Maria et al. (Hgg.): Sprechwirkung. Grundfragen, Methoden und Ergebnisse ihrer Erforschung. Berlin: Akademie Verlag, S. 9-58.

Stock, Eberhard/ Suttner, Jutta (1991): Wirkungen des Stimm- und Sprechausdrucks. In: Krech, Eva-Maria et al. (Hgg.), S. 59-142.

Trojan, Felix (1952): Der Ausdruck der Sprechstimme: Eine phonetische Lautstilistik. Wien u.a.: Verlag für Medizinische Wissenschaft Wilhelm Maudrich.

Zilliken, Franziska (1991): Beziehungskonstituierende Wirkungen des Sprechausdrucks. Eine empirische Untersuchung (Europäische Hochschulschriften 97). Frankfurt/Main u.a.: Peter Lang Verlag.

Christopher Nolan: The Dark Knight (USA 2008). DVD-Video, 146 Min. Bildformat 2.40:1/ 16:9, Ton: Deutsch/Englisch Dolby Digital 5.1. Warner Bros. Pictures.

### **Verwendete Internetseiten:**

„Internet Movie Database“. URL: <http://www.imdb.de/name/nm0005132/> (05.08.09).

„Jungstars.net – Martin Umbach Biographie“. URL: <http://www.jungstars.net/martin-umbach/vita.htm> (05.08.09).

„Mediacenter – tagesspiegel.de“. URL: <http://www.tagesspiegel.de/medien/hermes/cme1,226238.html> (05.08.09).

„Trailerseite.de – Batman: The Dark Knight Film Trailer“. URL: <http://www.trailerseite.de/archiv/trailer-2008/batman-the-dark-knight-trailer.html> (05.08.09).

„You Tube – The Dark Knight Synchro-Vergleich (Englisch/Deutsch)“. URL: <http://www.youtube.com/watch?gl=DE&hl=de&v=kW3KP3NgIBs> (21.07.09).



# Anhang

## **A. Für die Analyse verwendete Userkommentare auf „You Tube – The Dark Knight Synchro-Vergleich (Englisch/Deutsch)“**

<<http://www.youtube.com/watch?gl=DE&hl=de&v=kW3KP3NglBs>>

(abgefragt am 21.07.09)

TimoRac (vor 11 Monaten)

Jäger spielt sehr gut. Aber stimmlich kommt er einfach nicht a Ledger ran. Muss er sonst auch nicht.

Ein O-Ton Abklatsch muss nicht sein. Aber hier ist mir die Differenz einfach zu groß.

dummeFettehobbitse (vor 11 Monaten)

Ich finde es ist deutlich bemerkbar, dass Simon Jaeger sich anstraengt und im Grunde gar nicht schlecht ist, ihm fehlt jedoch die Tiefe der Stimme Ledgers.

ChrizzlieChrist (vor 11 Monaten)

Simon Jäger macht gute arbeit...um die Arbeit toll machen zu können hätte er sich wohl auch wie Ledger ein Monat in ein Hotelzimmer sperren müssen...kranke Joker Tagebücher schreiben und trainieren wie Puppenspieler sprechen...

Take2or3 (vor 11 Monaten)

Die deutsche Stimme ist perfekt. Ständig dieses Gemecker, eine tiefe Stimme passt nicht zu diesem Killer Clown! Von daher: Gute Arbeit!

patbiege (vor 11 Monaten)

Sorry, aber Perfekt ist was anderes Take2or3.

Das hat so wirklich nichts mit Ledger gemeinsam.

Ok Gut im gegensatz zu den Trailern hat er sich stark verbessert, aber Jäger wirkt verdammt unsicher, als wenn er nicht wüsste was er mit seiner Stimme machen soll. Schade, werd wohl doch nur OV ansehen können!

Take2or3 (vor 11 Monaten)

Ist doch wohl Ansichtssache. Im Vergleich zu Ledgers Stimme ziehe ich diese beim Joker vor. Es muss auch nichts mit der Original Stimme zu tun haben, diese kann sich auch so behaupten. Und ich bleibe dabei, für den Joker ist Jägers Stimme perfekt.

Ledgers Stimme für sich allein ist natürlich auch gut. War schließlich das erste, was man vom Joker gehört hat. Wichtig für mich ist jetzt nur noch, dass die Dialoge nicht zu weit abweichen. Ne tiefe Stimme können sie sich für Bane aufheben.

PrincessSalira (vor 11 Monaten)

Wow, der letzte Satz kommt im Deutschen fast noch verbitterter rüber als im Original! Man sollte überigens bedenken, dass Ledger hier schauspielerisch kaum Nachahmbares abliefern, Jäger stand/steht also vor einer riesigen Herausforderung -- die er bisher mit Bravour meistert.

BBGTet (vor 11 Monaten)

Für eine neue Synchro ist es jetzt bestimmt schon zu spät[wenn man bedenkt das der Film am 21.8. rauskommt].Wiedererkennungwert in allen Ehren aber für diese Roll past Simon einfach nicht...

BBGTet (vor 11 Monaten)

...[vllt hätte er wie Heath 2 Monate in einem Hotelzimmer hocken müssen um stimmlich zum Joker zu werden] aber dieses \*wir müssen jetzt schnell n Film synchronisieren Simon haste nochmal Zeit\* ist mit Abstand der größte Tiefpunkt in der SynchronGeschichte. Kennt jemand Kontaktdaten von Simon Jäger, dem Synchronstudio oder Warner Bros Germany?

vigeous (vor 11 Monaten)

Ich find's furchtbar. Jede Synchronisation. Sowas sollte nicht mehr erlaubt werden, höchstens gute Untertitel.

Schrecklich. Heath Ledger hat einfach keine hohe Stimme. Die Synchro hat noch nie zu ihm gepasst.

Friesa360 (vor 11 Monaten)

The German voice is cool.^ I think the german voice is a bit better for the Joker then the english one, because the German voice is not so deep. A higher voice is definitively better for the Joker.

Gasflasch (vor 11 Monaten)

schon komisch, bei der ersten Szene ist die deutsche Stimme dunkler^^

RealJoda89 (vor 11 Monaten)

Ich find das auch alles in Ordnung. Finde, dass das Irre auch auf Deutsch gut rüberkommt. Natürlich is original immer besser, aber was will man machen?

Take2or3 (vor 11 Monaten)

Also ich finde die deutsche Synchro außergewöhnlich gut, da sie nicht so dunkel und tief klingt. Erst war ich skeptisch, aber nun steht einem außergewöhnlichen Kinoereignis nichts mehr im Wege! =)  
Long live the Joker äh the Bat! ;)

LuckyLeny92 (vor 11 Monaten)

Jaa würd ich auch sagen... nicht gut, aber auch nicht total übertrieben scheiße. Trotzdem - eher schlecht als gut :-(  
Ich mag eigentlich die "deutsche Stimme" von Heath Ledger, aber ich finde, zu Joker passt sie nicht. Da müsste sie tiefer und furchteinflößender sein...

Ryker2007 (vor 10 Monaten)

Habe OV und Deutsche Fassung gesehen. Der Synchronsprecher macht seine Sache imho sehr gut, das Gänsehaut Feeling von Ledgers Originalstimme kommt natürlich nicht rüber. Die zahlreichen Oneliner die der Joker von sich gibt lassen sich ebend schwer ins Deutsche rübernehmen.

chaZz4479 (vor 10 Monaten)

Heaths original stimme ist viel besser und drastischer.  
das merkt man an der scene wo Heath im original "Hit me" ruft!  
das deutsche ist müll, genauso wie schon in Brokeback Mountain.  
schaut euch die movies in englisch an, und vergleicht..

voraussetzung ist dass ihr englisch einwandfrei versteht!

*chefkoch9* (vor 10 Monaten)

Finde die deutsche Stimme keinesfalls schlechter, im Gegenteil, sie bringt das bößer Clown Image irgendwie viel besser rüber als die originale.

Aber gut, manche müssen immer an was rummeckern was ihnen irgendwie nicht passt.

*jevnica* (vor 10 Monaten)

ich bin zweisprachig aufgewachsen und schaue nur originalfassungen, weil die deutschen versionen oft(immer) einfach zu schlecht und mit zu wenig aufwand produziert wurden.

ich hatte den film schon in england gesehen und fand die orginalstimme hatte zu wenig 'psycho'.

das einzige, was bei der deutschen version gefehlt hat waren die hintergrund geräusche. der synchronsprecher hat top arbeit geleistet, sowohl im vergleich mit der o-version, als auch mit anderen deutschen versionen irgendwelcher filme

*obscurita* (vor 10 Monaten)

Mr. Ledger hat mit dem Film ein unsterbliches machwerk hinterlassen. Der bis jetzt beste Joker im TV und Comic. Auch wenn ich ihn auf deutsch wegen der rotzigen Stimme mehr mag.Denn Schauspieler brauchen keine überragende Stimme. Nicht vergessen Schauspieler nicht Synchronisator. Egal gebt dem Mann einen Oscar, den hat er sich mehr als verdient. Alass my old friend.

*ninjaroooh* (vor 10 Monaten)

das ist mal ne gute deutsche synchronisation ;) der wahnsinn kommt gut raus.

*Just2Another4Psycho* (vor 10 Monaten)

Alle die meinen die deutsche Synchro sei besser als das englische Original haben sich TDK sicher nur auf Deutsch angeschaut.Sicher, die Synchro ist GUT, aber trotzdem KEIN Vergleich zum Original!Wie könnt ihr auch nur ein Original mit einer Synchro vergleichen??Bei der Synchro fehlt es an Tiefe, Leidenschaft.Das Original ist einfach..natürlicher, böser.Das kann selbst der beste Synchrosprecher nicht hinkriegen! Wer einmal TDK auf englisch gehört hat, wird den Joker nie wieder vergessen!

*jo5101* (vor 10 Monaten)

Auf deutsch wirkt sie psychopathischer und des passt zum Joker

*anarchyinthekitchen* (vor 9 Monaten)

ich find die deutsche noch wahnsinniger

*LanaPaluka84* (vor 9 Monaten)

habt recht ...in deutsch klingt er noch wahnsinniger

*nachtm4hr* (vor 8 Monaten)

ja ich wollte es grad sagen^^ also die deutsche synchronisation ist mal echt gelungen eigentlich... er klingt geil und viel irrer ^^

*Arrows* (vor 8 Monaten)

Sehr gute Synchro der krankhaften Stimme des Jokers. Hier hat man sich echt Mühe gegeben - klasse!

*darthschmidt* (vor 7 Monaten)

also es ist wirklich beides ausgezeichnet  
wobei der deutsche es so mehr ins psychopathisch humorische zieht  
aber beide ausgezeichnet

*intersheet* (vor 7 Monaten)

Meiner Erfahrung nach finden alle die die Synchro besser als das Original, die nicht gut genug Englisch können. Will niemandem vor den Kopf stoßen, ist aber meiner Erfahrung nach so. Ich habe den komplett auf Englisch und Deutsch im Kino gesehen und - obwohl die Synchro keineswegs schlecht ist - ist das Original meiner Meinung nach besser. Die Art wie er redet ist viel beängstigender. Aber wenn ich 50% dessen was er auf Engl sagt nicht verstünde, würde ich wohl auch die Deu besser finden.

*Noxilein2008* (vor 7 Monaten)

Die deutsche Synchronstimme ist ja mal absolut erstklassig! Da merkt man z.B. die Beeinflussung der Narben beim sprechen, worauf in der englischen Version kein bisschen geachtet wurde. Auch das plötzliche Umschwenken in einen aggressive Ton, in dieses raue, krächzende Stimme, wie es hier im Video beim letzten Vergleich gut zu hören ist, kommt im Original so gut wie nicht rüber und ist in der deutschen Version absolut genial! Sowas bringt erst die richtige Stimmung rüber!

*seW760i* (vor 6 Monaten)

Die Synchronisation ist den gesamten Film über genial!  
Das "Schmatzen" in der deutschen Version ist künstlerische Freiheit, würde ich sagen. Und das kommt irrsinnig gut rüber!!!  
Es macht den ohnehin schon grausam sarkastischen Joker noch verrückter und abstoßender^^  
Ich liebe den Joker :)

*Miyu707* (vor 6 Monaten)

meistens ist es eben so das die original stimme besser ist als die synchro. aber hier muss ich sagen, dass ich schon schlimmere synchros gehört hab, tatsächlich finde ich sie ganz okay...natürlich reicht die NICHT an Heagh's stimme ran, die einfach einzigartig genial gefährlich und verrückt ist!!! XD 5\*

*CrimsonThunder90* (vor 6 Monaten)

ich weiß gar nicht wieso alle meckern ich, find die synchro oscarreif xD die ist so... wahnsinnig^^

*Kinslaya* (vor 6 Monaten)

Finde die deutsche Synchro besser, weil man da dieses wahnsinnige "Schmatzen", das der Joker oft macht besser hört. Ich find das gibt ihm dieses Irre, Wahnsinnige noch mehr

*SeJens* (vor 6 Monaten)

also ich hab mir den film gestern auf dvd reingezogen und ich bin total begeistert von dem film. allein schon wie heath ledger den joker gespielt hat. aber das ist ja nicht grad das thema. also ich finde, die deutsche synchro passt viel besser dazu. allein schon wegen dem viel erwähnten schmatzen, dass den joker so verrückt macht.

*BaalCalom* (vor 6 Monaten)

auch wenn hier alles nur wiederholt wird...

ich muss den meisten hier zustimmen. durch dieses schmatzen und die merkwürdige stimme kommt er im deutschen noch wahnsinniger rüber, inhaltlich knapp unter dem niveau des originals.

Meisterspritze (vor 6 Monaten)

oh mann ey ich hab den film einmal auf deutsch gesehen und die synchro des films ist so zum erbrechen!!! Seit jahren noch nie sowas schlechtes gesehen!! und euer komisches schmatzen macht der auch im original nur mit so penetrant. durch die deutsche stimme wirkt der joker einfach zu albern!!! und das passt gar net

Makaay01 (vor 6 Monaten)

Ich finde, dass Simon Jäger den Joker sensationell gesprochen hat.

Durch das Schnalzen und Schmatzen kommt der Joker noch einen Taken verrückter rüber, was ja auch gut ist.

Auch die Lache, wo Joker diesen einen Batman anfängt zu töten, ist genial!

Ganz knapp nur unter dem Original, manchmal aber auch drüber.

youkilein (vor 5 Monaten)

Die deutsche Stimme klingt echt verrückter, manchmal cooler, die englische klingt gruseliger.

NeXx5 (vor 6 Monaten)

englisch is ind em defintiv cooler...

die englische stimme is tiefer udn rauher kommt eifanch geil als bösewicht...

die deutsche stimme find ich prinzipiell auch gut aber i wie passender als den guten fidn ich =)

wie s man von lleadger auch gewohnt is

Pastellela (vor 6 Monaten)

sicherlich ist die originalversion die beste (!), da kommt auch nie eine synchro ran, allein schon wegen der hintergrundgeräusche etc..

ich hab den film auf beiden sprachen gesehen und muss vielen zustimmen; auf deutsch klingt joker einfach weitaus wahnsinniger und mehr wie ein psycho. auf english eher normal, was nicht heißt, dass das schlecht ist.

ich finde, dass es da eher zwei unterschiedliche joker gibt. der eine wahnsinniger als der andere, aber jeder fabelhaft für sich.

RIP heath.

AgeOfLegacy77 (vor 6 Monaten)

Also ich persönlich finde bei manchen Stellen einfach die Deutsche und bei manchen Stellen die Originalversion besser. Trotzdem bin ich mehr für die Synchronstimme, sie klingt einfach verrückter und das ist beim Charakter und Aussehen vom Joker eines der wichtigsten dinge. Das Lachen bei der Szene mit Batman war auch genial, also großes Lob and Simon Jäger. @ Meisterspritze: Sei mal lieber froh das man in Deutschland überhaupt Synchronstimmen benutzt, schreib ne Bewerbung undmachsbesser!!

*oMystery4lifeOo* (vor 5 Monaten)

hallo?! die synchronstimme is doch absolut geil! die hört sich so richtig fies an^^ nya, die originalstimme is zwa immer noch die beste, aber im deutschen is se genau richtig. also wer hier sagt, die läge völlig daneben hört wohl nich mehr richtig!!!! also wirklich..

*KellerBeatsRecords* (vor 5 Monaten)

ich finde die deutsche stimme hat noch mer böshafte..so wie ich mir den Joker in den Comics vorstelle.. sogar noch besser

RIP Heath Ledger

wer wird woll der nächste joker sein?? könnte mir johnny depp sehr gut in der rolle vorstellen oder noch besser Mickey Rourke

*DaveMcIroy* (vor 5 Monaten)

Simon Jäger ist wirklich gut, aber Heaths Stimme ist rauer, brutaler, intensiver.

*ForeverCharmed87* (vor 5 Monaten)

Am besten ist sowieso immer das Original. Die Stimme passt zuHeath aber naja, O-Ton ist wie schon gesagt intensiver, rauer und viel aggressiver. Deswegen guck ich mir diesen Film auch nur auf Englisch an, weils besser ist und die Atmosphäre sowieso^^.

*koenigseggfan3* (vor 4 Monaten)

ich finde beide gleichgut ...aber auf deutsch kommt der joker iwie noch verrückter vor und das passt finde ich mehr

*jochenjacksparrowe1* (vor 3 Monaten)

ich weiß nicht, warum ihr -2 bei °english klingt besser° gebt das englische klingt besser

das deutsche passt nich zu dem °Joker°  
das verückte kommt so nich rüber

ich finde das englische auch viel besser

*sunshine836* (vor 3 Monaten)

boah ich liebe die deutsche synchronstimme von heath eindeutig is deutsch besser die stimme passt viel besser zu heath als seine echte stimme außerdem finde ich ihn mit der stimme viel verrückter.. aba egal ob joker oder net ich LIEBE heath

*SamKkDeluxe* (vor 3 Monaten)

aaaaaaaaaaaaah das ist auf englisch VIIIIIIIEEEEEEL besser !!!

schaut ma den ganzen film auf englisch, dann werdet ihr merken, dass der joker im original noch viel gruseliger rüberkommt !!!!!

*schwarzemamba42* (vor 2 Monaten)

ich finde, dass der joker im original viel fieser klingt im deutschen klingt seine stimme irgendwie schräg und ein bisschen ich nenns mal witzig

aber im grunde genommen kann man die originalsynchro nicht als deutschsprachler bewerten, da man sie ja ganz anders aufnimmt als englischsprachler

*mpw4c (vor 1 Woche)*

ich bin seehr zwiegespalten..

einerseits is englisch ein muss wegen heath ledgers stimme

andererseits klingt er auf deutsch irgendwie etwas gefährlicher..

ich kann nicht sagen was besser ist xD

mfg

*tennismile (vor 4 Tagen)*

auf deutsch kommt der viel abgedrehter rüber ...





## B. Daten der Akustischen Analyse

	Sprache	Video	Party	Vater (1. Teil)	Gordon	Batman	Hospital	Mittelwert
<b>Zeit (s)</b>	E	21,75	9,03	3,1	47,51	81,9	69,3	
	D	21,63	9,42	3,1	50,1	82	71	
<b>F0 (Hz)</b>	E	147,94	178,56	181,08	106,01			153,4
	D	190,79	184,71	190,67	128,17			173,6
<b>min. F0</b>	E	47,67	101,11	91,95	48,02			72,1
	D	45,94	75,52	92,94	47,39			65,4
<b>max. F0</b>	E	416,1	278,04	269,1	227,6			297,7
	D	399,27	307,81	280,95	397,22			346,3
<b>Range</b>	E	368,43	176,93	177,15	179,58			225,5
	D	353,33	232,29	188,01	349,83			280,9
<b>Intensität (dB)</b>	E	77,18	77,21	77,61	71,36			75,8
	D	75,97	77,22	76,77	74,27			76,1
<b>min. Intens.</b>	E	55	51	56	58			55
	D	53	59	56	56			56
<b>max. Intens.</b>	E	88	86	87	84			86,3
	D	85,5	86	87	81			85,6
<b>Pausenzeit</b>	E	5,1	2	0,8	17,5	26,2	16,5	
	D	3,5	1,1	0,9	14,5	25	11,3	
<b>Sprechzeit</b>	E	16,7	7	4	30	55,7	52,8	
	D	18,1	8,3	4,2	35,6	57	59,7	
<b>Artikulationsrate (Silben/Sek.)</b>	E	3,29	4,29	4	3,47	3,55	3,5	3,68
	D	3,81	3,98	5	3,43	4,28	3,77	4,05
<b>Pausenzeit - Differenz</b>		1,6 (31,4%)	0,9 (45 %)	0,1 (11,1%)	3 (17,1%)	1,2 (4,6%)	5,2 (31,5%)	