

# **PERSUASION UND PROSODIE**

---

UNTERSUCHUNG ZUR PERZEPTION  
EMOTIONALER SPRECHWEISEN  
AM BEISPIEL EINER PARFUMWERBUNG.

**Dissertation**  
**zur Erlangung des akademischen Grades**  
**Doctor philosophiae (Dr. phil.)**

**vorgelegt dem Rat der Philosophischen Fakultät**  
**der Friedrich-Schiller-Universität Jena**

von Beate Redecker (M.A.)  
geboren am 13. Januar 1976 in Stollberg/Erzgeb.

1. Gutachter: Prof. Dr. phil. Adrian P. Simpson
2. Gutachter: Priv.-Doz. Dr. phil. Baldur Neuber

Tag des Kolloquiums: 7. Dezember 2006

**Für Christoph.**

# INHALTSVERZEICHNIS

Verzeichnis der Abbildungen	v
Verzeichnis der Tabellen	vii

## **I Theoretische Grundlagen und Stand der Forschung**

<b>1</b>	<b>EINLEITUNG</b> .....	1
<b>2</b>	<b>PERSUASION UND PROSODIE</b> .....	4
2.1	Persuasion: Begriffsbestimmung .....	4
2.2	Das Modell der Elaborationswahrscheinlichkeit (ELM elaboration likelihood model) .....	6
2.3	Prosodie: Begriffsbestimmung .....	14
2.4	Persuasion und Prosodie: Forschungsüberblick .....	18
2.4.1	Sprechgeschwindigkeit .....	21
2.4.2	Grundfrequenz .....	23
2.4.3	Lautstärke .....	25
2.5	Zusammenfassung .....	26
<b>3</b>	<b>EMOTION UND PROSODIE</b> .....	28
3.1	Emotionskonzepte .....	28
3.2	Vokaler Emotionsausdruck und Prosodie: Das Linsenmodell von Brunswik in einer Adaption von Scherer .....	34
3.3	Emotion und Prosodie: Forschungsüberblick .....	38
3.3.1	„Push“ und „Pull“ Faktoren .....	40
3.3.2	Sprecherorientierte Arbeiten: „Encoding studies“ und „Transmission studies“ .....	43
3.3.3	Hörerorientierte Arbeiten: „Decoding studies“ .....	48
3.4	Zusammenfassung .....	52

---

<b>4</b>	<b>REZEPTIONSDETERMINANTEN UND PROSODIE</b> .....	55
4.1	Externe Rezeptionsdeterminante: Sprechausdruck .....	55
4.2	Interne Rezeptionsdeterminante: Hörgewohnheit .....	67
4.3	Lautsymbolismus und Ikonizität .....	70
4.4	Zusammenfassung .....	74
<b>II</b>	<b>EMPIRISCHER TEIL</b>	
<b>5</b>	<b>EXPERIMENTELLES DESIGN</b> .....	75
5.1	Fragestellung und Zielsetzung .....	75
5.2	Material und Methoden .....	77
<b>6</b>	<b>ERGEBNISSE</b> .....	86
6.1	Apparative Analyse .....	86
6.2	Auditive Analyse .....	111
6.3	Wirkungsanalyse .....	121
6.4	Zusammenfassung der Ergebnisse .....	135
<b>7</b>	<b>DISKUSSION</b> .....	137
<b>8</b>	<b>ZUSAMMENFASSUNG UND AUSBLICK</b> .....	149
<b>9</b>	<b>LITERATUR</b> .....	151
<b>10</b>	<b>DANKSAGUNG</b> .....	167
<b>11</b>	<b>LEBENS LAUF</b> .....	168
<b>12</b>	<b>EHRENWÖRTLICHE ERKLÄRUNG</b> .....	170
<b>A</b>	<b>ANHANG</b> .....	171
A.1	Fragebogen zur Wirkungsanalyse .....	172
A.2	Fragebogen zur auditiven Analyse .....	175
A.3	Statistik (Rohdaten) .....	177

## VERZEICHNIS DER ABBILDUNGEN

<b>Abb. 1:</b>	Einstellungsänderung als Funktion von Ablenkung und Qualität der Argumente	11
<b>Abb. 2:</b>	Das ELM von Petty & Cacioppo (1986) im Überblick	13
<b>Abb. 3:</b>	Schematische Darstellung der Dichotomie der Prosodie	16
<b>Abb. 4:</b>	Zweidimensionale Darstellung von Emotionsbezeichnungen: vertikale Dimension = Aktivität, horizontale Dimension = Valenz	31
<b>Abb. 5:</b>	Brunswiks Linsenmodell in einer Adaption von Scherer	35
<b>Abb. 6:</b>	Inferenzprozess am Beispiel von Extrovertiertheit	37
<b>Abb. 7:</b>	Das Brunswiksche Modell in einer Adaption Scherers zur Systematisierung der Forschungslage	39
<b>Abb. 8:</b>	Zusammenhang zwischen der Dichotomie des Prosodiebegriffs und einer Systematik zur Einordnung der Literatur	54
<b>Abb. 9:</b>	Strichfiguren als Grundlage für Lautsymbole	72
<b>Abb. 10:</b>	Schematische Darstellung des experimentellen Designs	77
<b>Abb. 11:</b>	Bildfolgen und Filmsequenzen des Werbespots	80
<b>Abb. 12:</b>	Lautsymbolik des Namens des Parfums und des Slogans	82
<b>Abb. 13:</b>	Temporale Struktur aller drei Sprechfassungen im Vergleich	87
<b>Abb. 14:</b>	Temporale Struktur und Lautsegmentierung von „Caractère“ aller drei Sprechfassungen im Vergleich	90
<b>Abb. 15:</b>	Temporale Struktur und Lautsegmentierung von „Macht am Mann“ aller drei Sprechfassungen im Vergleich	94
<b>Abb. 16:</b>	Temporale Struktur und Lautsegmentierung von „das Markante am Mann“ aller drei Sprechfassungen im Vergleich	96
<b>Abb. 17:</b>	Grundfrequenzverlauf von „Caractère“ aller drei Sprechfassungen im Vergleich	101
<b>Abb. 18:</b>	Darstellung des melodischen Verlaufs von „Caractère“ aller drei Sprechfassungen in Halbtonschritten	102
<b>Abb. 19:</b>	Darstellung des melodischen Verlaufs von „Macht am Mann“ aller drei Sprechfassungen in Halbtonschritten	104
<b>Abb. 20:</b>	Grundfrequenzverlauf von „Macht am Mann“ aller drei Sprechfassungen im Vergleich	105

---

<b>Abb. 21:</b>	Grundfrequenzverlauf von „das Markante am Mann“ aller drei Sprechfassungen im Vergleich	108
<b>Abb. 22:</b>	Darstellung des melodischen Verlaufs von „das Markante am Mann“ aller drei Sprechfassungen in Halbtonschritten	109
<b>Abb. 23:</b>	Profildarstellung der Stimmqualitätsmerkmale aller drei Sprechfassungen im Vergleich	113
<b>Abb. 24:</b>	Profildarstellung der Verlaufskriterien aller drei Sprechfassungen im Vergleich	118
<b>Abb. 25:</b>	Graphische Darstellung der Altersverteilung und der Geschlechterverteilung innerhalb der Stichprobe	122
<b>Abb. 26:</b>	Zusammensetzung der Stichprobe bezüglich der regionalen Herkunft	122
<b>Abb. 27:</b>	Graphische Darstellung der Mittelwerte und Standardfehler für die positiven Merkmale des Werbespots	124
<b>Abb. 28:</b>	Graphische Darstellung der Mittelwerte und Standardfehler für die negativen Merkmale des Werbespots	125
<b>Abb. 29:</b>	Graphische Darstellung der Mittelwerte und Standardfehler für die positiven Merkmale der Person	127
<b>Abb. 30:</b>	Graphische Darstellung der Mittelwerte und Standardfehler für die negativen Merkmale der Person	128
<b>Abb. 31:</b>	Graphische Darstellung der Mittelwerte und Standardfehler für die Dufteigenschaften von Faktor ‚lieblich-fruchtig‘	129
<b>Abb. 32:</b>	Graphische Darstellung der Mittelwerte und Standardfehler für die Dufteigenschaften von Faktor ‚frisch-vitalisierend-sportlich‘	131
<b>Abb. 33:</b>	Graphische Darstellung der Mittelwerte und Standardfehler für die Dufteigenschaften von Faktor ‚würzig-holzig-herb‘	132
<b>Abb. 34:</b>	Graphische Darstellung des Zusammenhangs zwischen dem Geschlecht und dem Kaufverhalten (Chi-Quadrat-Test nach Pearson).	133
<b>Abb. 35:</b>	Graphische Darstellung des Zusammenhangs zwischen dem Kaufverhalten und den drei Sprechfassungen (Chi-Quadrat-Test nach Pearson).	134

## VERZEICHNIS DER TABELLEN

<b>Tab. 1:</b>	Zwei Modelle aus der Werbepsychologie im Vergleich: AIDA-Formel und PPPP-Prinzip	5
<b>Tab. 2:</b>	Prosodische Sprachsignalkomponenten	18
<b>Tab. 3:</b>	Scherers Komponenten-Prozess-Modell (SECs)	33
<b>Tab. 4:</b>	Zusammenfassung stimmlicher Emotionseffekte im Vergleich zu neutraler Sprechweise	47
<b>Tab. 5:</b>	Vergleich der Prozentzahlen für das Diskriminieren einzelner Emotionen in zwei empirischen Studien	50
<b>Tab. 6:</b>	Parameter der Stimme und Sprechweise nach Fährmann.	59
<b>Tab. 7:</b>	Mittlere Sprechstimmlage für die unterschiedlichen Stimmgattungen nach Wirth	61
<b>Tab. 8:</b>	Kriterienkatalog für die auditive Analyse des Sprechmaterials	83
<b>Tab. 9:</b>	Kategorien des Fragebogens	84
<b>Tab. 10:</b>	Stichprobe der Fragebogenuntersuchung	85
<b>Tab. 11:</b>	Sprechgeschwindigkeit innerhalb der einzelnen Sprechsequenzen aller drei Sprechfassungen im Vergleich	89
<b>Tab. 12:</b>	Temporale Struktur von „Caractère“ in den drei Sprechfassungen im Vergleich	91
<b>Tab. 13:</b>	Temporale Struktur von „Macht am Mann“ in den drei Sprechfassungen im Vergleich	93
<b>Tab. 14:</b>	Temporale Struktur von „das Markante am Mann“ in den drei Sprechfassungen im Vergleich	98
<b>Tab. 15:</b>	Melodieverlauf von „Caractère“ in den drei Sprechfassungen im Vergleich	103
<b>Tab. 16:</b>	Melodieverlauf von „Macht am Mann“ in den drei Sprechfassungen im Vergleich	107
<b>Tab. 17:</b>	Melodieverlauf von „das Markante am Mann“ in den drei Sprechfassungen im Vergleich	110
<b>Tab. 18:</b>	Übersicht über die Mittelwerte und Standardabweichungen der auditiven Analyse	120
<b>Tab. 19:</b>	Darstellung der mittleren Differenz, der Standardabweichung und der daraus resultierenden Signifikanzen zwischen den einzelnen Gruppen für die Wirkung aller Items des gesamten Werbespots im Vergleich	126

<b>Tab. 20:</b>	Überblick über distinktive Eigenschaften der Stimme und Sprechweise der präferierten Sprechfassung (SF II).	135
<b>Tab. 21:</b>	Zusammenhang zwischen den Kategorien des Fragebogens und den Emotionsdimensionen.	140

# 1 EINLEITUNG

Die Auseinandersetzung mit prosodischen Phänomenen ist schon früh Gegenstand wissenschaftlicher Untersuchungen gewesen und bietet auch aktuell immer wieder Anlass zu neuen Fragestellungen. Die Interessengebiete sind dabei im Wesentlichen in drei inhaltliche Schwerpunkte unterteilbar: (a) Arbeiten, die sich mit der Perzeption von sprecherimmanenten Gestaltungsweisen und der Identifikation von individuumsspezifischen Ausdrucksmustern beschäftigen, (b) Arbeiten, die das Decodieren und Encodieren von Emotionen durch stimmliche und sprecherische Merkmale in den Blickpunkt ihrer Betrachtungen rücken und (c) Arbeiten, die im Zusammenhang mit Sprachsyntheseverfahren den Einfluss prosodischer Stimuli fokussieren. Geleitet sind die Untersuchungen maßgeblich durch die dichotome Betrachtung prosodischer Phänomene unter formellen und funktionellen Aspekten.

In der vorliegenden Arbeit wird diese Dichotomie weiter untersucht und insbesondere der Frage nachgegangen, inwieweit prosodische Merkmale die Persuasion und die Perzeption im Kontext einer Parfumwerbung beeinflussen können. Den ersten Teil und zugleich den theoretischen Rahmen der Arbeit bilden Kapitel 2, 3 und 4, in welchen Form-Funktions-Beziehungen der Prosodie in drei Teilgebieten unter Berücksichtigung des aktuellen Standes der Forschung nachgegangen wird.

In Kapitel 2 wird dazu ein renommiertes Modell (Elaboration Likelihood Model von Petty & Cacioppo 1986) als Bezugsrahmen zur Einordnung von Persuasionseffekten vorgestellt und im Zusammenhang mit Parametern der Stimme und Sprechweise diskutiert. Anhand verschiedener Arbeiten wird zudem aufgezeigt, dass stimmliche und sprecherische Merkmale in wesentlichen Anteilen die Erzeugung von persuasiven Effekten konstituieren.

Kapitel 3 beleuchtet Komponenten des vokalen Emotionsausdrucks in Interaktion mit prosodischen Parametern. In diesem Zusammenhang wird das Linsenmodell von Brunswik (1956) in einer Adaption von Scherer (1978, 2000a) als theoretisches Fundament zum experimentellen Design der Arbeit vorgestellt. Dieser Ansatz integriert neben phonetischen auch sprechwissenschaftliche Methoden der Datengewinnung und stellt somit eine sichere Basis für die Dateninterpretation dar.

In Kapitel 4 werden ausgehend vom empirischen Teil der Arbeit ausgewählte prosodische Parameter definiert und systematisiert, um sie auf diese Weise in auswertbare Kategorien zu bringen. Darüber hinaus werden für wichtige prosodische Parameter die Möglichkeiten und die Grenzen der Datengewinnung und der Dateninterpretation von apparativen und auditiven Verfahren diskutiert.

Aufbauend auf diesen theoretischen Rahmen wird im zweiten Teil der Arbeit, in Kapitel 5 und 6 ein Experiment beschrieben, was der Frage nach der Perzeption und Persuasion emotionaler Sprechweisen im Kontext einer Parfumwerbung nachgeht.

Kapitel 5 beinhaltet die exakte Beschreibung des methodischen Designs. In einer Fragebogenuntersuchung wurden 156 Probanden gebeten, ein Werbevideo in seiner gesamten Wirkung und bezüglich der Wirkung der dargestellten Person und des erwarteten Duftes zu beurteilen. Dabei wurden drei separaten Hörergruppen drei unterschiedliche Fassungen des Werbespots präsentiert, die sich ausschließlich in prosodischen Merkmalen der Stimme und Sprechweise unterschieden. Die Probanden wurden im Vorfeld der Befragung in keiner Weise für stimmliche Phänomene sensibilisiert. Ein systematischer Fehler durch eine Voreingenommenheit und gelenkte Aufmerksamkeit der Versuchspersonen sollte so ausgeschlossen werden (Blindstudie). Die drei Sprechfassungen wurden darüber hinaus einer apparatephonetischen Untersuchung und einer auditiven Untersuchung in Form eines Expertenurteils unterzogen, um so die distinktiven Merkmale der einzelnen Sprechfassungen herauszuarbeiten.

In den Ergebnissen der Untersuchung, die in Kapitel 6 beschrieben sind, kann gezeigt werden, dass bereits kleinste stimmliche und sprecherische Modifikationen, eingebettet in den Kontext zusätzlicher visueller und akustischer Stimuli des Werbespots, signifikante Veränderungen in der Persuasion (betrifft die Kaufeinstellung) und in der Perzeption (betrifft die Wahrnehmung) des Werbespots nach sich ziehen. Dies umfasst neben der Wahrnehmung des gesamten Werbespots auch die Perzeption einzelner Kategorien, die allein durch ausgewählte prosodische Merkmale als signifikant positiver oder negativer beurteilt werden. Interessanterweise bleibt die Darstellung des assoziierten Duftes von prosodischen Merkmalen völlig unberührt.

Darüber hinaus werden in einer Wirkungsuntersuchung Merkmale der Stimme und Sprechweise extrahiert, die Hörerüberindividuell präferiert und somit auch

als besonders wirkungsvoll beurteilt werden. Dazu zählen eine sonore Bassstimme, die im Durchschnitt ca. eine kleine Terz unter den in der Literatur angegebenen Werten liegt, der Einsatz von Modalstimme ohne begleitendes Hauchen und Knarren, keine Registerwechsel, dezente Melodieintervalle um ca. eine große Terz oder Quarte herum und ein hoher Anteil an Stimmhaftigkeit. In der sich anschließenden Diskussion in Kapitel 7 wird aufgezeigt, wie sich die Ergebnisse des Experiments in die bestehende Forschungsliteratur einordnen lassen und welche Ansatzpunkte sich für nachfolgende Untersuchungen ergeben.

## 2            **PERSUASION UND PROSODIE**

### 2.1            **PERSUASION: BEGRIFFSBESTIMMUNG**

In der traditionellen Rhetorik ist die Persuasion der Effekt von rhetorischen Fertigkeiten und bezeichnet Kommunikationsprozesse, die nicht primär der Informationsübermittlung dienen. Der gewünschte Kommunikationseffekt kann dabei vornehmlich auf emotionaler Basis im Sinne von *überreden* oder auf rational-argumentativer Basis im Sinne von *überzeugen* erzielt werden.

Beim Überreden wird das Bezugssystem des Hörers kurzgeschlossen, meist mit emotionalem Überdruck und der Suggestion, hier verwirkliche sich seine Wunschwelt, wird er zum Handeln im Reflex gebracht; zum Handeln mit Reflexion dagegen im Zusammenwirken seiner rationalen, emotionalen und voluntativen Kräfte, beim Überzeugen. (Geißner 1969:54f)

Die Verwendung des Terminus <Kommunikationseffekt> deutet an, dass sich beide Formen im Ergebnis der Wirkungsdauer unterscheiden, wobei *überzeugen* als resistenter in der Zugänglichkeit für Gegenargumente beurteilt wird als *überreden* (Roß 1994:203). Die Kunstfertigkeit der Persuasion liegt in der geschickten Art und Weise, neue Informationen mit bereits vorhandenen zu verknüpfen, um somit von Bekanntem auf Neues, Unbekanntes, zu schließen. Persuasion i.w.S. meint damit

[...] den gesamten vom Kommunikator initiierten Überzeugungsvorgang mit dem Ziel, bei anderen einen Standpunktwechsel herbeizuführen. (Knappe 2003:874)

Auf diesen persuasiven Wechsel bezieht sich wohl auch Cicero, wenn er von der Rhetorik als einer „seelenwendende[n] Königin“ spricht (Knappe 2003:874). Die Voraussetzung für die Auslösung persuasiver Prozesse ist die mentale Differenz zwischen den Kommunikationspartnern, denn die Ursache für Persuasion liegt in einer bestimmten Intention des Kommunikators, einen Wechsel von Urteilen, Einschätzungen oder Haltungen unter verschiedenen Aspekten beim Gegenüber herbeizuführen. Knappe (2003:875) führt folgende Orientierungsaspekte an, um so einen Wechsel herzustellen: (a) instruktiv (etwas ist so und nicht anders), (b) verifikativ (wahr/falsch; wahrscheinlich/unwahrscheinlich), (c) evaluativ (gut/schlecht; schön/hässlich), (d) axiomativ (gilt/gilt nicht), (e) emotiv (lieben/hassen) und (f) voluntativ (tun/nicht tun). Dass in der Werbe-

branche von dieser Form der Meinungs- und Haltungsbeeinflussung reger Gebrauch gemacht wird, liegt nahe.

Aufbauend auf diese Aspekte haben sich in der Werbepsychologie Modelle herausgebildet, wie z.B. die AIDA-Formel oder das PPPP-Prinzip, die solche persuasiven Elemente als systematischen Bestandteil in ihre Inhalte integrieren. Diese Modelle, die der amerikanischen Werbepsychologie entstammen, finden in der einschlägigen Literatur (z.B. Frindte 2001:153f, Felser 2001:15f, Kloß 2000:78f) große Berücksichtigung, weswegen sie stellvertretend in dieser Arbeit angeführt werden sollen, obwohl es eine Vielzahl weiterer Modelle zur Beschreibung der kommunikativen Werbewirksamkeit gibt. Ihnen ist als wesentliches Element gemein, dass sie den Rezipienten zum Handeln anregen sollen. Unklar bleibt allerdings, ob es sich dabei um deskriptive oder um präskriptive Modelle handelt. Mit anderen Worten: Wird beschrieben, wie Werbung wirkt oder wie Werbung wirken soll (Felser 2001:15).

Bei der AIDA-Formel stehen die einzelnen Buchstaben für die jeweiligen Schritte, die durch die Werbebotschaft beim Rezipienten ausgelöst werden sollen (mit dem Ziel, das Produkt zu (ver)kaufen). Bei PPPP handelt es sich um ein Prinzip, wodurch beim Rezipienten, ausgehend von einem bildlichen Impuls (picture), die Aufmerksamkeit auf das Produkt gelenkt werden soll.

AIDA-Formel			PPPP-Prinzip		
<b>A</b>	Attention	Aufmerksamkeit erregen	<b>P</b>	Picture	Bildliche Darstellung
<b>I</b>	Interest	Interesse wecken	<b>P</b>	Promise	Versprechen
<b>D</b>	Desire	Wunsch wecken, etwas haben zu wollen	<b>P</b>	Prove	Beweisführung
<b>A</b>	Action	Handlung auslösen (kaufen)	<b>P</b>	Push	Anstoß zum Handeln

**Tab. 1: Zwei Modelle aus der Werbepsychologie im Vergleich: AIDA-Formel und PPPP-Prinzip.**

Strukturell gesehen laufen persuasive Kommunikationsprozesse immer darauf hinaus, einen Wechsel herbeizuführen, aber nicht jeder Wechsel muss unbedingt eine Verhaltensänderung nach sich ziehen. Müller (1999:13) differenziert

in Anlehnung an Hovland et al. (1966) je nach zeitlichem Wirkungseinfluss drei verschiedene Wechsel und beschreibt diese als: (a) *Opinion change* (Meinungswechsel): kurzfristiger interner Wechsel, (b) *Attitude change* (Einstellungswechsel): langfristiger interner Wechsel und (c) *Behaviour change* (Handlungswechsel): langfristiger Wechsel, der extern durch die Handlung der Person sichtbar wird. Werbekampagnen können je nach Idee (Produkt, Zeit, Budget, Zielgruppe ...) alle drei Formen des Wechsels zum Ziel haben. Dabei entscheidet der beabsichtigte Wechsel über das Ziel strategisch-kommunikativer Praxis, um eine gewünschte Reaktion auf eine kalkulierte rhetorische Handlung hervorzurufen.

Für diese Arbeit ist nicht der Wirkungsmechanismus der verbalargumentativen Struktur von Interesse, sondern der Einfluss von persuasiven Merkmalen, die die Argumentation begleiten, so wie es z.B. die prosodischen Faktoren tun. Dazu soll im folgenden Abschnitt ein renommiertes Modell vorgestellt werden, was eine Grundlage für die Einordnung prosodischer Merkmale in den Persuasionsprozess darstellt. Es bildet zugleich den theoretischen Rahmen für die Darstellung prosodischer Merkmale als persuasive Elemente im Kommunikationsprozess.

## 2.2 DAS MODELL DER ELABORATIONSWAHRSCHEINLICHKEIT (ELM ELABORATION LIKELIHOOD MODEL) VON PETTY & CACIOPPO

Nothing is so practical as a good theory.  
Kurt Lewin<sup>1</sup>

Mitte der achtziger Jahre entstand das ELM (elaboration likelihood model) von Petty & Cacioppo (1986). Es stellt neben anderen Modellen aus dieser Zeit, z.B. dem AAP – Attribution Analysis of Persuasion von Eagly et al. (1981) oder dem HSM – Heuristic-Systematic Model von Chaiken et al. (1989) den Versuch dar, einen theoretischen Rahmen für Persuasionseffekte zu erstellen. Dabei bildet es einen guten Bezugsrahmen für eine allgemeine Theorie der Einstellungsänderung, weil es in seiner Konzeption verschiedene Elemente wie theoretische Aspekte zur Einstellungsänderung und empirische Ergebnisse integriert. Das Modell versucht unterschiedliche z.T. widersprüchliche Befunde über

---

<sup>1</sup> zitiert nach Hovland et al. (1966:2)

persuasive Kommunikation und Einstellungsänderung zu beschreiben und bildet somit einen umfassenden Rahmen für das Verständnis grundlegender Prozesse bezüglich der Wirksamkeit persuasiver Kommunikation (Güttler, 1996:192). Für Petty & Cacioppo ist Persuasion die Einstellungsänderung durch Kommunikation, eine Definition, die für diese Arbeit einen relevanten theoretischen Hintergrund bildet.

Der empirisch gestützte Ansatz des ELM (s.u.) wird von Petty & Cacioppo in der Monographie „Communication and Persuasion. Central and Peripheral Routes to Attitude Change“ zusammengefasst. Darin postulieren die Autoren Richtlinien, die gleichzeitig als Grundlage für ihre Theorie der Einstellungsänderung herangeführt werden können (Petty 1986:5, Übersetzung aus: Güttler, 1996:192f, Frindte, 2001:87f):

- Menschen sind motiviert, an adäquaten Einstellungen festzuhalten. Die Angemessenheit der Einstellung wird durch interpersonale Vergleiche gewonnen.
- Die Gründlichkeit der Elaboration einer Botschaft variiert mit individuellen und situativen Faktoren.
- Die Elaboration von Argumenten kann relativ objektiv (datengesteuert) oder relativ verzerrt [biased] (voreingenommen) erfolgen.
- Die Motivation und die Fähigkeit zur Informationsverarbeitung sind die zentralen Variablen, von denen die objektive oder die verzerrte Verarbeitung abhängt.
- Wenn die Motivation und die Fähigkeit, Argumente zu verarbeiten abnehmen, werden periphere Hinweisreize zu bedeutenden Determinanten der Persuasion.

Das ELM unterscheidet zwei prototypische Modi der Einstellungsänderung, die sich antagonistisch gegenüberstehen: eine zentrale Route (central route) und eine periphere Route (peripheral route). Diese Differenzierung bildet das Kernstück des Modells und soll im folgenden Abschnitt skizziert werden.

Bei der zentralen Route handelt es sich um das kritische Abwägen und die kognitive Auseinandersetzung mit der argumentativen Struktur einer Botschaft im Kommunikationsprozess. Darunter versteht man z.B. die Investition von Zeit, die sorgfältige Prüfung der Argumente, das Einbringen von Wissen, die tiefe gedankliche Elaboration, die Integration der Argumente in vorhandene kognitive Schemata und die Selbstgenerierung von themenrelevanten und sachbezogenen Argumenten.

Die periphere Route hingegen ist durch eine Vielfalt an Mechanismen charakterisiert, die das Verbalargumentative und damit den Ideengehalt einer Botschaft begleiten wie: die soziale Identifikation, die Konditionierung oder die Verwendung von Heuristiken. Es handelt sich dabei um Persuasionsprozesse, die nicht auf tiefgründigem, themenrelevantem Denken beruhen, sondern die nichtinhaltliche Merkmale wie Glaubwürdigkeit, Sympathie, Sachverstand, Attraktivität und andere Hinweise (cues) aus der Kommunikation in den Mittelpunkt rücken. An dieser Stelle gewinnt das ELM eine zentrale Bedeutung für die hier vorliegende Arbeit, denn neben Merkmalen wie Musik, Bildern, Farben u.a. sind es maßgeblich die prosodischen Merkmale eines Sprechers (Stimme und Sprechweise), die in der Fernsehwerbung die Persuasion der peripheren Merkmale (Sympathie, Attraktivität, Glaubwürdigkeit ...) provozieren.

Menschen verfügen nur über begrenzte Zeit und Ressourcen und können deshalb nicht jede persuasive Botschaft, der sie ausgesetzt sind, bis ins Detail verarbeiten. Daher stellen die peripheren Prozesse oftmals den entscheidenden Impuls im Wirkungsmechanismus dar. Die Art des Persuasionspfades ist nach Petty & Cacioppo von der Motivation und von der Fähigkeit des Empfängers abhängig, die vorgegebene Botschaft zu verarbeiten. Motivation und Fähigkeit sind die Variablen, die die Elaborationswahrscheinlichkeit bestimmen. Je größer diese Wahrscheinlichkeit ist, desto größer ist auch die Verarbeitung über die zentrale Route. Die Einstellungen, die über diesen Weg gewonnen werden, gelten als andauernder und widerstandsfähiger. Sind die Empfänger weniger motiviert oder besitzen sie nur wenig Wissen über ein Thema, dann lassen sie sich weniger durch die Argumente als vielmehr durch die die Argumentation begleitenden Elemente beeinflussen. Diese Aussagen leiten die Autoren von verschiedenen Experimenten ab, in denen der Zusammenhang von der Qualität der Argumente und der persönlichen Relevanz untersucht wird (Petty & Cacioppo 1986:141ff).

Gorn (1982) untersuchte beispielsweise die Wirkung von angenehmer Musik in Werbespots und konnte aufzeigen, dass in Abhängigkeit von der persönlichen Relevanz der Werbung, Musik unterschiedlich starke persuasive Wirkungen nach sich zieht. Probanden, die eine enge Beziehung zu dem beworbenen Produkt haben, entschieden sich zu 71% für den informationsbasierten Werbespot, wohingegen sich Probanden mit niedriger persönlicher Relevanz zu 63% für das Produkt, welches mit angenehmer Musik beworben wurde, entschie-

den. Diese Parallele zwischen dem Grad von Motivation und der Wirkung von Hintergrundmusik benennen auch Park & Young (1986) in einem ähnlichen Experiment. Sie argumentieren aber dahingehend, dass Hintergrundmusik Leute mit hoher Motivation von argumentativen Prozessen abhält. Eine Untersuchung von Chebat et al. (2000) geht der Bedeutung von Persuasion in einem abgewandelten Kontext nach, mit der Frage danach, ob Hintergrundmusik in einem Geschäft Käufer positiv beeinflusst. In ihrer Untersuchung finden die Autoren heraus:

Music does not moderate significantly the effects of the salespersons [...] but low and moderately arousing music [...] does influence significantly the effects on the acceptance of the salesperson's arguments and the 'desire to affiliate,' i.e., to enter into communication. (Chebat et al. 2000:405)

Güttler (1996:194) nennt für die primär periphere Verarbeitung von Botschaften die Benutzung von einfachen heuristischen Regeln und die Verwendung einfacher Schemata wie:

- Experten sagen immer die Wahrheit.
- Sympathischen Leuten kann man vertrauen.
- Fachleute wissen es am besten.
- Was teuer ist, muss auch gut sein.
- Was ich gekauft habe, muss ich mögen.

An diese Form der Informationsaufnahme und Informationsverarbeitung seitens der Rezipienten knüpfen Werbefachleute an, wenn sie sich weniger inhaltlichen Argumentationen als vielmehr der Verwendung solcher kognitiven Schemata zuwenden, um ihre Produkte oder Personen attraktiv und sympathisch, glaubwürdig und kompetent erscheinen zu lassen. Die Frage nach Attraktivität spielt in der Werbepsychologie eine entscheidende Rolle. Bower (2001) untersuchte die Attraktivität von Models und fand dabei heraus, dass Frauen, die Werbung anschauen, die in Werbespots dargestellten „highly attractive models“ ablehnen, und dass dies darüber hinaus mit einer negativen Beurteilung der Effektivität des Sprechers einhergeht.

The pattern of results from two studies indicates that, when sufficient negative affect is generated as a consequence of comparison with beautiful models, evaluations of both the model as a spokesperson and the product argument may be affected adversely because of model derogation. (Bower 2001:51)

Und in einer Untersuchung von Reichert et al. (2001) zeigen die Autoren auf, inwieweit Sexappeal von Werbedarstellern Einfluss auf kognitive Prozesse und Persuasionsprozesse hat. Sie fassen zusammen:

[...] sexual appeals were more persuasive overall than matched nonsexual appeals for social marketing topics<sup>2</sup>. Sexual appeals also stimulated more favorable ad executionrelated thoughts but had a negative effect on cognitive elaboration (e.g., support and counterarguments). (Reichert et al. 2001:13)

Außerdem scheint die Art der Darbietung einer Botschaft (mündlich oder schriftlich) einen maßgeblichen Einfluss auf den Persuasionsweg zu haben. Müller (1999:36) schreibt:

[...] u.a. bekennen Petty & Cacioppo schließlich, daß im Falle mündlicher Kommunikation periphere Merkmale die Persuasion in weit stärkerem Maße beeinflussen als bei schriftlicher Kommunikation. [...] Daß stark involvierte bzw. motivierte Empfänger durch die ‚central route‘, d.h. vorrangig durch Argumente überzeugt werden, muß damit auf den schriftlichen Kontext beschränkt werden.

Wie lässt sich nun erklären, dass die periphere Route in mündlichen Situationen einen stärkeren Einfluss auf die Persuasion hat? Petty & Cacioppo (1981) erklären es damit, dass die peripheren Merkmale simultan mit Argumenten verarbeitet werden und es somit zu einer Konfrontation zwischen den Argumenten und den peripheren Merkmalen kommt. Dies führt zu einer Ablenkung von den Argumenten und lässt die peripheren Merkmale als die eigentlichen Persuasionsdeterminanten in mündlichen Redesituationen in den Vordergrund treten (Müller 1996:36f).

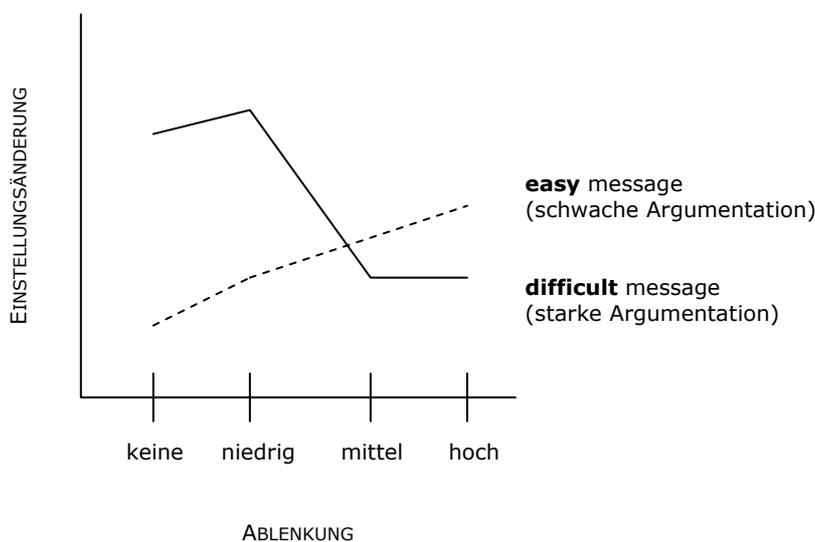
Das Motiv der Ablenkung wurde in der Forschung verschiedentlich untersucht, wobei unterschiedliche Ergebnisse aus den einzelnen Studien hervorgehen. Mehrere Untersuchungen haben gezeigt, dass Ablenkung den Persuasionserfolg reduziert (Bohner 2003:284). Doch Petty et al. (1976) fanden in zwei Experimenten ein differenzierteres Ergebnis heraus. In einem Versuch hörten die Vpn eine dargebotene Botschaft zum Thema Erhöhung der Studiengebühren über Kopfhörer. Die zwei persuasiven Botschaften (je drei Minuten) enthielten in einem Fall fünf triftige (starke) Argumente und im anderen Fall fünf faden-scheinige (schwache) Schlüsselargumente. Parallel dazu erschien den Vpn in Abhängigkeit zum Grad der Ablenkung auf einem Monitor in unterschiedlichen

---

<sup>2</sup> Social marketing definieren die Autoren in Anlehnung an Fine (1992:3) als „the application of commercial marketing methods to help create demand for ‚social products‘: [for example] energy conservation, women’s rights, 55-mph speed limit, cancer research, gay rights, seat belt use, [ect.]“

Frequenzen<sup>3</sup> ein „x“. Sie sollten dann jeweils eine Angabe zu dem Quadranten machen, in welchem das „x“ erschien. Im Ergebnis halten die Autoren fest, dass das Motiv der Ablenkung mit der Produktion von Gegenargumenten in Zusammenhang steht. In Abbildung 1 sind die Ergebnisse zusammengefasst: mit zunehmender Ablenkung wird bei schwacher Argumentation das Hervorbringen von Gegenargumenten gesteigert, wohingegen bei schwer zu verstehenden Botschaften die befürwortenden Gedanken (favorables thoughts) abnehmen.

A Message X Distraction interaction indicated that distraction tended to enhance persuasion for the counterarguable message but reduce persuasion for the message that elicited primarily favorable thoughts. (Petty et al. 1976:874)



**Abb. 1: Einstellungsänderung als Funktion von Ablenkung und Qualität der Argumente. (Petty et al. 1976:879)**

Nach dem ELM wird angenommen, dass Persuasionsvariablen verschiedene Funktionen erfüllen können, wobei Personen unterschiedlich anfällig für die Verarbeitung von zentraler und peripherer Route sind. Die Art des Persuasionsweges wird dabei nicht als ein „entweder-oder“, sondern vielmehr als ein Kontinuum verstanden. Als Beispiel führt Bohner (2003:283) ein attraktives Model an, welches für ein Shampoo wirbt. Folgende Einstellungen sind denkbar:

<sup>3</sup> für niedrige Ablenkung 4 „x“ pro Minute, für mittlere Ablenkung 12 „x“ pro Minute und für hohe Ablenkung 20 „x“ pro Minute.

- peripherer Hinweisreiz (die Elaborationswahrscheinlichkeit ist gering) der Persuasionserfolg wird unmittelbar über periphere Mechanismen vergrößert (Konditionierung)
- Model als Argument einer Botschaft (die Elaborationswahrscheinlichkeit ist hoch) der Persuasionserfolg wird über thematisch relevante kognitive Reaktionen vergrößert (Wenn ich das Shampoo benutze, werden meine Haare so schön aussehen wie ihre.)
- Attraktivität des Models verstärkt die Motivation, die Botschaft zentral zu verarbeiten (Sieh dir das attraktive Model an, ich sollte genau darauf achten, was sie sagt.)

Das ELM ist in der einschlägigen Fachliteratur vielseitig beschrieben und diskutiert und bietet auch für diese Arbeit einen guten Aufhänger. Ein gelegentlich in der Literatur angeführter Kritikpunkt ist allerdings, dass Petty & Cacioppo zwar zugestehen, dass zentrale und periphere Prozesse gleichzeitig ablaufen, dass sie aber keine Mechanismen und Bedingungen über deren Zusammenspiel angeben.

Diese Arbeit wird sich im Fokus mit der Frage beschäftigen, inwieweit prosodische Merkmale einen Einfluss auf die Perzeption eines Werbespots haben und inwieweit sie dadurch auch die Persuasion (Einstellung zum Produkt bzw. zum Kauf des Produkts) nachhaltig verändern können. Sämtliche verbalargumentativen Impulse bleiben unberücksichtigt und werden aus der Untersuchung ausgeklammert.

In Abbildung 2 soll das Modell überblicksartig dargestellt werden (Petty & Cacioppo 1986:4, Güttler 1996:196f), wodurch die wichtigsten Stadien der persuasiven Kommunikation und deren Zusammenwirken transparent gemacht werden sollen. Die drei möglichen Endpunkte (a) zentrale positive oder negative Einstellungsänderung, (b) periphere, temporäre Einstellungsverschiebung (...shift) und (c) keine Änderung der Einstellung werden dabei hervorgehoben. Die zentrale Route entspricht dem Hauptstrang. Vier Bedingungen des Hauptstranges münden in den Punkt <periphere Hinweisreize vorhanden> und bilden die periphere Route, die eine vorläufige Einstellungsverschiebung nach sich ziehen kann, wenn periphere Hinweisreize vorhanden sind. Die Einstellungsverschiebung fließt dann wieder in den Hauptstrang zurück und kann dort über die zentrale Route zu einer (positiven oder negativen) Einstellungsänderung mit hoher Elaborationswahrscheinlichkeit führen.

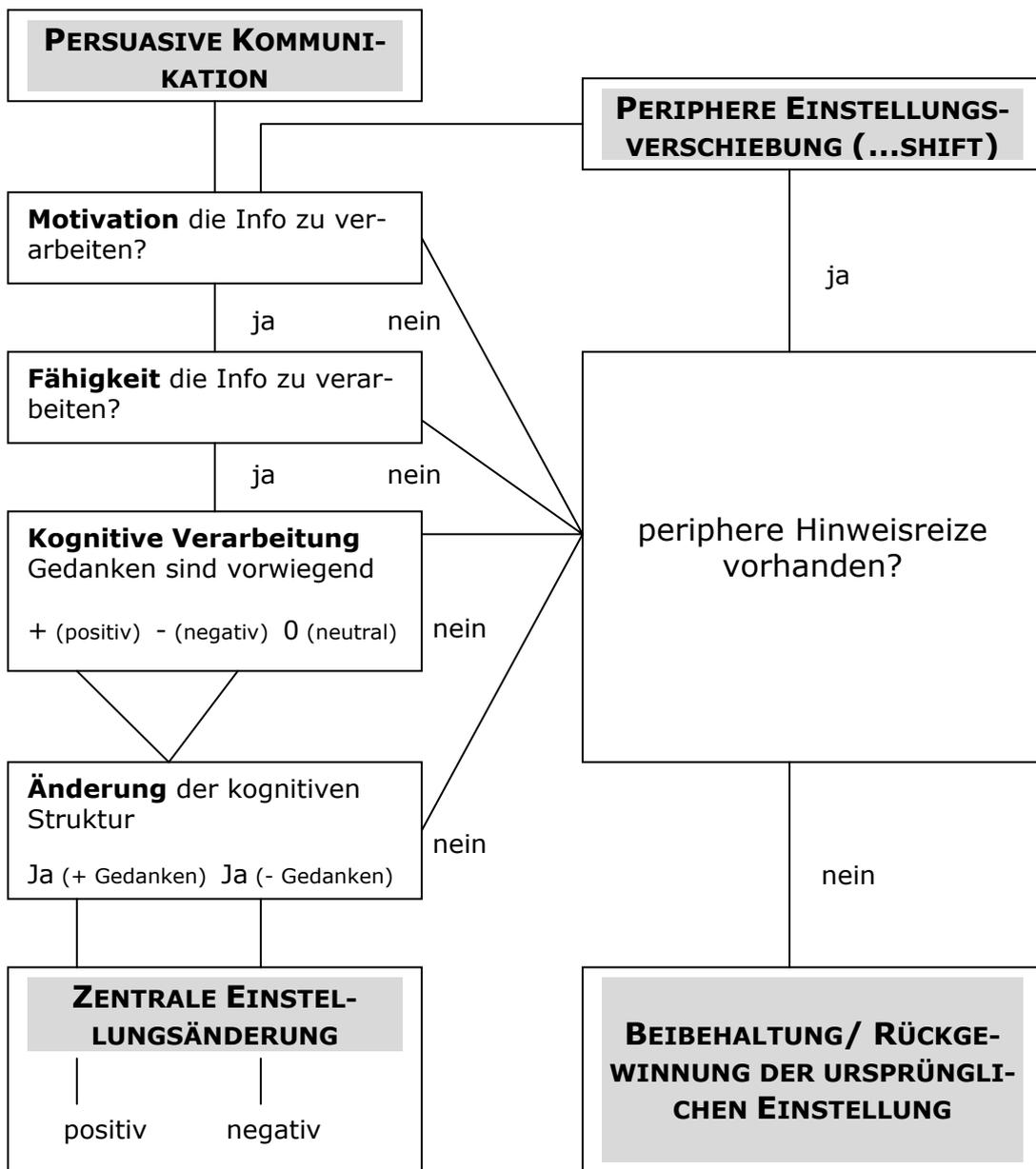


Abb. 2: Das ELM von Petty & Cacioppo (1986) im Überblick.

## 2.3 PROSODIE: BEGRIFFSBESTIMMUNG

In diesem Kapitel soll der Terminus <Prosodie> in seiner Verwendung für die vorliegende Arbeit definiert werden, um Missverständnissen wie sie in der einschlägigen Literatur nicht selten vorliegen, vorzubeugen, die sich aus einer unscharfen und unpräzisen Verwendung des Begriffs ergeben. Der Begriff leitet sich von dem griechischen *prosōdía* ab, was soviel wie „Zugesang“ bedeutet, und setzt sich aus den Bestandteilen *ōidē* <Gesang> und *prós* <hinzu> zusammen. Demnach bezeichnet er Elemente, die zu den Segmenten hinzukommen, sie begleiten, unterstützen, hervorheben oder konterkarieren. Welche Elemente zur Prosodie hinzuzählen und welche Gewichtung einzelne Elemente erfahren, wird in verschiedenen Schulen und Strömungen unterschiedlich verstanden, woraus sich breite terminologische Streuungen ergeben.

Die Begriffe Prosodie, Intonation und Suprasegmentalia sind in der Fachliteratur divers besetzt. Dies ist nicht zuletzt auf die Verwendung der Begriffe in unterschiedlichen Fachbereichen zurückzuführen, was besonders in der Phonetik, Linguistik und Psychologie zu unterschiedlichen Konnotationen führt (Neuber 2002:15). Aber auch innerhalb der Phonetik herrscht Uneinigkeit über die begriffliche Präzision und die Abgrenzung der Begriffe voneinander. In linguistischen und sprachwissenschaftlichen Wörterbüchern werden die drei Begriffe nahezu synonym verwendet. So findet man bei Bußmann (1990:352f, 618) unter dem Eintrag *Prosodie*:

[...] Gesamtheit sprachlicher Eigenschaften wie Akzent, Intonation, Quantität [...]. Sie beziehen sich [...] auf Einheiten, die größer sind als ein einzelnes Phonem.

Und unter dem Eintrag *Intonation* schreibt er:

[...] Gesamtheit der prosodischen Eigenschaften von sprachlichen Äußerungen, die nicht an den Einzellaut gebunden sind.

Eine zweite Problematik resultiert aus der definitorischen Divergenz eines engen und eines weiten Intonationsbegriffes. So umfasst bei Glück (2000:315) Intonation i.w.S.

[...] die Gesamtheit der prosodischen Eigenschaften [...] lautsprachlicher Äußerungen, d.h. Akzent- und Pausenstruktur [...] sowie des Tonhöhenverlaufs [...].

Als Vertreter dieser Auffassung können angeführt werden (Neuber 2002:16): v. Essen (1956), Heike (1969), Crystal (1975), Pheby (1975) und Bolinger (1989). Intonation i.e.S. bezeichnet hingegen nur den

[...] Stimmtönenverlauf [...] unter dem Aspekt der Satzmelodie [...]. Verlauf der Grundfrequenz [...]

Nöth (1991:23) nennt für diese Auffassung als Vertreter: Bannert (1985), Kohler (1977), Pierrehumbert (1981) und Wunderlich (1988).

In dieser Arbeit werden die Begriffe Prosodie, Intonation und Suprasegmentalia in Anlehnung an Neuber (2002) und Kranich (2003) synonym verstanden. Im Folgenden wird aus konventionellen Gründen nur noch der Begriff der Prosodie verwendet.

Es sollen in dieser Arbeit nicht die verschiedenen Definitionen, die in der Literatur kontrovers diskutiert werden, vorgestellt werden (vgl. dazu Neuber 2002:15ff), vielmehr soll ein Aspekt, den Neuber in seinen Ausführungen stark macht, Berücksichtigung finden, nämlich die Frage nach der Dichotomie der Prosodie in einen formalen und einen funktionalen Aspekt. Dazu schlägt er zwei Definitionen vor: eine (a) parameterorientierte, akustische Definition (formaler Aspekt), unter der der Autor die Zusammenfassung der

[...] auditiv wahrnehmbaren Merkmale Melodieführung, Lautheit und Lautheitsdifferenz, Akzent, Sprechgeschwindigkeit (Tempo) und Tempowechsel, Sprechrhythmus, Pausen, (indexikalisch bedingte) Stimmqualität, Stimmausdruck (Timbre) und Gesamtdauer der prosodischen Erscheinung

versteht (Neuber 2002:51f). Und des Weiteren eine (b) funktionale Definition, unter der der Autor in Anlehnung an Stock (1996a) Prosodie als

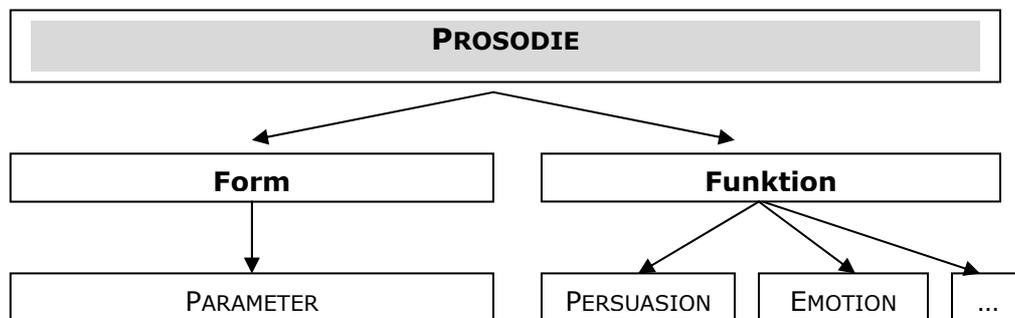
[...] ein von vielen Faktoren bestimmtes, begrenzt autonomes multifunktionales System beschr[eibt] [...], das in Konkurrenz mit anderen Kommunikationsmitteln fungiert. [...] im Sinne Bühlers zugleich als Symptom und Signal.

Ergänzend fügt Neuber hinzu, dass Prosodie unter bestimmten Umständen auch (begrenzte) Symbolfunktion ausüben kann, die im Deutschen im konnotativen Bereich liegt (Neuber 2002:53).

Aus Stocks Ausführungen ergibt sich ein anderer Punkt, der in der Literatur kontrovers diskutiert wird, die Frage danach, ob Prosodie als ein autonomes Signalisierungssystem fungiert oder ob sie immer in Beziehung zu grammatischen Kategorien betrachtet werden muss (Selting 1995:16f). Als Vertreter der ersten Annahme, die die Prosodie als ein eigenständiges Signalisierungs-

system verstehen sind zu nennen: Bolinger (1986, 1989), Gibbon (1988) und Knowles (1984). Dem gegenüber stehen Intonationstheorien, die die Prosodie in systematischer Beziehung zu grammatischen Kategorien beschreiben. Hierfür seien Vertreter angeführt wie: Halliday (1967, 1995), Pheby (1980) und Wunderlich (1988). Selting schlägt in ihren Abhandlungen zur Prosodie (1995:16) eine Verbindung aus beiden Auffassungen vor, indem sie die Prosodie als ein autonomes, von Grammatik und Textsorte unabhängiges Signalisierungssystem, welches in Konkurrenz mit grammatischen und textuellen Signalen steht, definiert. Damit bezieht sie einen ähnlichen Standpunkt wie Stock (s.o.).

Neubers dichotome Beschreibung des Prosodiebegriffs und die damit einhergehende Trennung in eine Formseite und eine Funktionsseite ermöglicht die Darstellung unterschiedlicher Kategorien (Abb. 3). Mit den Fragen danach, welche Funktion prosodische Parameter in persuasiven Kontexten übernehmen können und welche Bedeutung prosodischen Parametern im Markieren von emotionalen Zuständen zukommt, werden zentrale Funktionsbereiche akustischer Parameter benannt, die gleichzeitig einen wichtigen theoretischen Rahmen für den empirischen Teil dieser Arbeit bilden.



**Abb. 3: Schematische Darstellung der Dichotomie der Prosodie.**

Obwohl Form und Funktion eng aneinander gebunden sind und in einem reziproken Verhältnis zueinander stehen, scheint eine differenzierte Betrachtung der beiden Aspekte aus strukturellen Gründen sinnvoll. Die prosodischen Eigenschaften, die ihren Fokus in den Formalia haben, ermöglichen z.B. die Analyse sprecherbezogener Merkmale des Sprechausdrucks, wie es im Kapitel 3 ‚Emotion und Prosodie‘ ausführlicher beschrieben wird. In diesem Kapitel soll

das Augenmerk auf den Zusammenhang von Persuasion und Prosodie gelenkt werden. Dazu scheint es sinnvoll, zunächst aufzuzeigen, welche Funktionsauffassungen von Prosodie in der Literatur vertreten sind und auf welcher Basis sie gewonnen werden. Neuber (2002:34ff) hat folgende Erkenntnismethoden zusammengefasst und systematisiert:

- theoretisch gewonnene Erkenntnisse (Löhken 1997)
- experimentelle Untersuchungen, die sich einzelnen Parametern widmen (Tischer 1993)
- Ableitungen der Intonationstheorien anderer Sprachen (Pheby 1980)
- empirische Untersuchungen authentischer Kommunikationsereignisse (Hollmach 1996)

Ergänzen lassen sich darüber hinaus Kombinationen aus den aufgezählten Herangehensweisen wie:

- annähernd phänomenologische Beschreibung (Pompino-Marschall 1995)
- Formbeschreibung mit syntaktischen und semantischen Interpretationen (Stock 1980)
- Form-Funktionsbeschreibungen mit dem Ziel didaktisierbarer Regeln (Stock 1996)
- konversationsanalytische Form-Funktionsbeschreibung (Selting 1995)
- sprachpsychologische Form-Funktionsbeschreibung (Scherer 1982)

Auf der Basis dieser Erkenntnismethoden lassen sich verschiedene prosodische Funktionen benennen. Ackermann et al. (1993:242) schlagen eine Dreiteilung der prosodischen Sprachsignalkomponenten vor, um prosodische Funktionen zu spezifizieren, die in Tabelle 2 dargestellt ist.

Dieser Einteilung kann man entnehmen, dass man sich unter wissenschaftlich sehr unterschiedlichen Fragestellungen den prosodischen Funktionen nähern kann. Während die linguistisch orientierte Prosodieforschung die prosodischen Komponenten in ein lexikalisch, syntaktisch eingebundenes Regelwerk integriert, versucht die pragmatische Prosodieforschung eher Auskunft über die Intention eines Sprechers zu geben, mit der er die Aufmerksamkeit des Zuhörers steuern will, ohne dabei vordergründig den lexikalisch-syntaktischen Rahmen zu berücksichtigen. Die Inhalte der affektiven Prosodieforschung konzentrieren sich auf die Beschreibung von vokalen, nonverbalen Botschaften der Stimme und Sprechweise und untersuchen primär expressives Lautgebungsverhalten,

welches die Persönlichkeitsmerkmale eines Sprechers wie Alter, Geschlecht, Affekt ... sowie dessen Emotionalität in den Mittelpunkt der Betrachtung rückt (vgl. Kapitel 3 Emotion und Prosodie).

LINGUISTISCHE PROSODIE	PRAGMATISCHE PROSODIE	AFFEKTIVE PROSODIE (expressives Lautgebungsverhalten)
<ul style="list-style-type: none"> <li>▪ Wortakzent (<i>umfahren/umfahren</i>)</li> <li>▪ Satzakzent, Satzfokus (das hast <i>Du</i> gesagt)</li> <li>▪ Satzmodus (Du hast gefragt ?, !, .)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Sprecherintention (Ironie)</li> <li>▪ Dialogsteuerung (Wortübergabe)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Persönlichkeit (Dominanz)</li> <li>▪ Emotionen, Affekte (Traurigkeit)</li> </ul>

**Tab. 2: Prosodische Sprachsignalkomponenten (Ackermann et al. 1993:242)**

Für die hier vorliegende Arbeit scheint die affektive Funktion der Prosodie im Vordergrund der Betrachtung zu stehen. Um die Wirkung von Stimme und Sprechweise jedoch hinreichend zu beschreiben, können linguistische und pragmatische Parameter nicht völlig vernachlässigt werden. Zudem lassen sich die unterschiedlichen Funktionsbereiche nicht immer klar voneinander abtrennen. So ist beispielsweise Ironie, welche in Tabelle 2 der pragmatischen Prosodie zugeordnet wird, in anderen Arbeiten (Reinke 1998) als Emotion und damit als Bestandteil der affektiven Prosodie beschrieben.

Im folgenden Abschnitt soll nun der Zusammenhang zwischen Persuasion und Prosodie vor allem vor dem Hintergrund funktionaler Aspekte der Prosodie genauer beschrieben werden.

## 2.4 PERSUASION UND PROSODIE: FORSCHUNGSÜBERBLICK

In wissenschaftlichen Abhandlungen stellt die Verbindung von Persuasion und Prosodie bisher keinen zentralen Forschungsschwerpunkt dar. Die Forschungslage gestaltet sich als eher schwierig, was u.a. daran liegt, dass beide Felder in separaten Wissenschaftsdisziplinen (Rhetorik, Linguistik, Phonetik, Sprechwissenschaft und Psychologie) erforscht werden, wobei die Frage danach, welche Funktion prosodischen Merkmalen im Persuasionsprozess zukommt, wei-

testgehend unbeachtet bleibt<sup>4</sup>. Sprechwissenschaftliche Untersuchungen, die sich mit der Sprechwirkungsforschung beschäftigt haben, bieten zwar einen guten Einblick in Grundfragen und Methoden, gehen aber darüber hinaus nicht konkret auf die persuasive Wirkung der Prosodie ein. Sie referieren vielmehr Hörerurteile, die u.a. Hinweise auf die Persuasion geben können, wie z.B. sympathisch vs. unsympathisch, leidenschaftlich vs. leidenschaftslos, einfühlsam vs. distanziert (Krech et al. 1991:229), sie können dabei aber das Zusammenwirken von Persuasion und Prosodie weder belegen noch widerlegen. Ab den 70er-Jahren entstand in Amerika eine Reihe von Publikationen, die sich mit einzelnen Parametern der Stimme und Sprechweise und deren Funktion in persuasiven Kommunikationssituationen beschäftigt haben (z.B. Pearce & Brommel 1972, Hall 1980). Inwieweit einzelne Parameter eine generalisierende Ableitung auf die Gesamtheit aller prosodischer Parameter erlauben, oder ob einzelne Parameter stärker als andere in den Wirkungsprozess integriert sind, kann an dieser Stelle nicht geklärt werden. Ebenso muss angemerkt werden, dass es bisher wissenschaftlich nicht geklärt ist, inwieweit sich die Ergebnisse aus der angelsächsischen Prosodieforschung auf das Deutsche übertragen lassen. Jüngere Arbeiten machen den interkulturellen Vergleich von Emotionsrezeption zu ihrem Inhalt und zeigen auf, dass es kulturelle Unterschiede in der Produktion und Perzeption einzelner Emotionen gibt, die auf bestimmte modellhafte Hörmuster zurückzuführen sind (Reinke 1998:191).

Wenn man die Methoden der empirischen Datengewinnung vergleicht, so lassen sich zwei Linien unterscheiden: (a) Studien, die ihre Ergebnisse deskriptiv gewinnen (dies ist meistens der Fall, wenn mehrere Parameter untersucht werden) und (b) Studien, die ihre Daten manipulativ erzeugen. Die Manipulation kann dabei durch natürliche Variationen oder auf synthetische Weise herbeigeführt werden. Arbeiten, die nur einen oder zwei Parameter prüfen, haben oftmals manipulativen Charakter.

Im Vergleich zu Arbeiten, die den Zusammenhang von prosodischen Merkmalen und Emotionalität untersuchen, gestaltet sich das empirische Design zur Untersuchung des Zusammenhanges von Persuasion und prosodischen Merkmalen als eher schwierig. Persuasion muss funktionieren und kann daher nicht

---

<sup>4</sup> Bsp. hierfür wären: Selting, M. (1995): *Prosodie im Gespräch*, oder Pompino-Marschall, B. (1990): *Die Silbenprosodie*. Beide Arbeiten nähern sich der Prosodie unter Fragestellungen, bei denen die Wirkung prosodischer Merkmale unberücksichtigt bleibt.

wie Emotionalität professionell reproduziert werden. Auch kann man eine Wirkung nicht direkt abfragen, sondern muss vielmehr auf indirektem Weg Daten gewinnen, die mögliche Rückschlüsse auf die Wirkung zulassen. Möglicherweise ist diese methodische Schwierigkeit der Datengewinnung und auch der Dateninterpretation ein Grund für den großen, noch ausstehenden Forschungsbedarf. Im empirischen Teil dieser Arbeit (Kap. 5 und Kap. 6) wird dieses Dilemma noch einmal an einem konkreten Experiment diskutiert.

Es sollen nun relevante Ergebnisse älterer und neuerer Arbeiten vorgestellt werden, die zugleich einen Überblick über die aktuelle Forschungslage bilden.

Eine jüngere Arbeit von zentraler Bedeutung stellt die Studie von Gélinas-Chebat et al. (1996) dar. In einem manipulativ erzeugten Sprechkorpus untersuchen die Autoren eine Kombination aus je zwei verschiedenen Graden an Lautstärke und Tonhöhenvariation [THV] (laut, leise, viel THV, wenig THV). Auf diese Weise ergeben sich 4 (2x2) phonetische Varianten, die wiederum mit zwei weiteren Graden an Empfängermotivation<sup>5</sup> verbunden werden, so dass ein Muster von 2x2x2 entsteht. Die zwei phonetischen Bedingungen werden in ihren Varianten von einem Berufsschauspieler auf zwei verschiedene Radiowerbungen angewendet. Die Werbung wird dabei so gestaltet, dass sie einmal von hoher und einmal von niedriger persönlicher Relevanz für den Empfänger ist (ELM s.o.), wobei die Texte linguistisch so ähnlich wie möglich gehalten werden. Die acht sich ergebenden Fassungen werden in unterschiedlichen Gruppen von Studenten (acht Gruppen mit insgesamt 279 Probanden) vorgespielt; diese sollen im Anschluss den Werbespot und den Sprecher beurteilen. Parallel dazu werden die Sprechfassungen akustisch analysiert. Die Ergebnisse zeigen, dass periphere Stimuli wie die Prosodie in Abhängigkeit von der Motivation des Empfängers maßgeblich zum Persuasionserfolg beitragen.

Intensity of voice affects credibility of the source significantly more under low than high involvement; intonation of voice affects credibility more under high than low involvement. (Gélinas-Chebat et al. 1996:243)

Zu ähnlichen Ergebnissen kommen Peterson et al. (1995). Hier werden in einem deskriptiven Verfahren 21 amerikanische Verkäufer (erfolgreiche und weniger erfolgreiche) auf Tonband aufgenommen. Der linguistische Aspekt wird nahezu konstant gehalten, da die Verkäufer ein vorformuliertes Verkaufsskript

---

<sup>5</sup> Der Aspekt der Empfängermotivation bildet einen direkten Bezug zum ELM, welches von der Annahme ausgeht, dass Motivation und Fähigkeit des Rezipienten die Persuasion beeinflussen.

erhielten. Die Aufnahmen werden 26 Hausfrauen<sup>6</sup> (sie stellen die größte Zielgruppe des beworbenen Produktes dar) vorgespielt, die die Sprecher in Form von semantischen Differenzialfragebögen beurteilen sollen. Die Bögen werden im Anschluss hinsichtlich der Perzeption der unterschiedlichen stimmlichen Merkmale ausgewertet. Dabei werden die Bewertungen der Wirkung des Sprechers mit der Haltung des Empfängers in Bezug gesetzt. Die auditive Bewertung wird durch eine akustische Analyse unterstützt. Aus den Ergebnissen geht hervor, dass Parameter wie Sprechgeschwindigkeit, mittlere Pausenlänge und Grundfrequenzverlauf signifikant mit dem Verkaufserfolg korrelieren. Im Vergleich zu einer Kontrollgruppe, die aus 13 jungen Männern im ähnlichen Alter besteht, die keine Verkäufer sind, ergeben sich vor allem Differenzen im Sprechtempo (Verkäufer sprechen schneller) und im Grundfrequenzverlauf (die Verkäufer produzierten „ansprechendere“ Muster). Die Autoren kommen daher zu folgendem Fazit:

They [...] reinforce prior conclusions regarding the importance of voice characteristics in the communication process and illustrate how subtly these characteristics influence interpersonal perceptions. From a personal selling perspective, the primary implication of the findings is clear: *how* a sales message is communicated may be as important as *what* is communicated with respect to output sales performance. (Peterson et al. 1995:12)

Wie oben erwähnt, wurde in verschiedenen Studien der Zusammenhang von einzelnen Parametern und der persuasiven Funktion, die damit einhergeht, untersucht. Dabei muss kritisch angemerkt werden, dass das Außerachtlassen anderer Parameter und vor allem deren Zusammenwirken die Gefahr einer sehr einseitigen Betrachtungsweise birgt. Im Folgenden sollen ausgewählte Arbeiten in ihren Grundaussagen und Ergebnissen vorgestellt werden, die sich mit den Parametern Sprechgeschwindigkeit, Grundfrequenz und Lautstärke und mit daraus resultierenden Persuasionserfolgen beschäftigt haben.

### 2.4.1 SPRECHGESCHWINDIGKEIT

Miller et al. (1976) belegen, dass Veränderungen des Sprechtempos allein keine Auswirkungen auf den Informationsprozess in der Kommunikationssituation

---

<sup>6</sup> Hier ergeben sich möglicherweise kulturelle Unterschiede zwischen den USA und Deutschland.

ausüben, dass aber schnelleres Sprechen die Glaubwürdigkeit und Kompetenz eines Sprechers erhöht.

That is, it might provide cues for viewing a fast speaker as more knowledgeable, competent, and facile. In this case a rapid speaking rate constitutes a credibility cue and on this basis should likewise enhance persuasion. (Miller et al. 1976:616)

Das Kriterium der Sprechgeschwindigkeit funktioniert somit als zentraler Marker in der Beurteilung der Kompetenz und Glaubwürdigkeit eines Sprechers. Ähnliche Ergebnisse erhielten Apple et al. (1979) in ihren Untersuchungen, bei denen sie die Sprechgeschwindigkeit eines Textes um 30% zu beiden Poolen variierten. Dabei veränderte sich die Wahrnehmung der sprechenden Personen deutlich. Langsame Sprecher werden als weniger ehrlich, weniger gewandt und passiver wahrgenommen.

[...] slow-talking men are judged to be less truthful, fluent, emphatic, serious and persuasive, and more passive, although they are also seen as more potent. (Apple et al. 1979:724)

In drei Experimenten können sie nachweisen, dass die akustischen Eigenschaften einer Äußerung einen erheblichen Einfluss auf die Beurteilung einer Person in Bezug auf deren Gemütszustand und deren Charakterzüge haben. Steward & Ryan (1982) betonen in ihren Ausführungen, dass schneller sprechende Personen als wohlgesonnener und dem Gegenüber gewogener eingeschätzt werden. Und Ergebnisse von Brown et al. (1985) belegen, dass „faster speakers“ als kompetenter gegenüber „slower speakers“ eingeschätzt werden. Zwei weitere Studien, die an dieser Stelle angeführt werden sollen, von Smith et al. (1975) und Woodall & Burgoon (1983), betonen zusammenfassend einmal mehr die enge Beziehung von Sprechtempo und sprecherimmanenten Wesenszügen. Demnach vergrößert ein erhöhtes Sprechtempo die persuasiven Effekte und zieht eher Charakteristika wie Glaubwürdigkeit, Objektivität und Überzeugungsgabe nach sich.

Those who talk fast, whether they are car salespersons, characters in television and radio commercials, or friends engaged in a heated discussion, often seem to have a persuasive advantage: their fast rate of speaking seems to enhance our impressions of them as well as the impact of their persuasive arguments. Above and beyond our own day-to-day experience, recent empirical evidence suggests that fast talking does indeed enhance source credibility and persuasion. (Woodall & Burgoon 1983:126)

## 2.4.2 GRUNDFREQUENZ

Eine vielseitig untersuchte Stimmeigenschaft ist die Grundfrequenz. In verschiedenen Studien werden ihr Verlauf und ihre Variabilität beschrieben. Hier sollen einige Arbeiten angeführt werden, die speziell Eigenschaften der Grundfrequenz im Zusammenhang mit der Wirkung eines Sprechers in den Mittelpunkt ihrer Betrachtung rücken. Es muss allerdings angemerkt werden, dass die Produktion und Perzeption von Grundfrequenzverläufen stark kulturabhängig ist, was vor allem Hirschfeld (1998:75f) in ihrer Arbeit thematisiert. Deshalb können die Ergebnisse angelsächsischer Studien nicht unkritisch und bedenkenlos auf das Deutsche übertragen werden.

Aus Angaben in der Literatur über Grundfrequenzbereiche, in denen Männer- und Frauenstimmen eine optimale Wirkung nach sich ziehen, ergibt sich bisher ein eher kontroverses Bild. Scherer (1979:186) beschreibt höhere Stimmlagen bei Männern als Zeichen für Extrovertiertheit, Entschiedenheit und Selbstvertrauen, solange sich der Stimmbereich nicht dem Bereich der Frauen nähert, welcher durchschnittlich eine Oktave höher liegt. Eckert & Laver (1994:37) widersprechen diesem Ergebnis und beschreiben eine gegenteilige Tendenz, wonach tiefe Männerstimmen als angenehm und als Ausdruck von Souveränität, Vertrauenswürdigkeit und Autorität eingeschätzt werden.

Oft sind es auch Klischeevorstellungen, die zur Bevorzugung dieser Stimmeigenschaft führen: Der Weihnachtsmann wird in der Regel mit einer tiefen Stimme dargestellt und ebenso Zeus oder der liebe Gott, wenn er in Film und Theater zu Worte kommt. [...] So ist es nicht verwunderlich, daß tiefe Männerstimmen sehr häufig am Ende eines Reklamespots für die zusammenfassende Bewertung des Produktes eingesetzt werden. (Eckert & Laver 1994:37)

Apple et al. (1979) führten drei Experimente durch, bei denen Probanden Tonbandaufnahmen vorgespielt bekamen, in denen Sprecher mit normaler Grundfrequenz zu hören waren, oder Sprecher, bei denen die Grundfrequenz um 20% nach oben bzw. unten manipuliert wurde. Dabei fanden sie heraus:

Speakers with high-pitched voices were judged less truthful, less emphatic, less 'potent' (smaller, thinner, faster), and more nervous. (Apple et al. 1979:715)

Burgoon et al. (1990) postulieren in Anlehnung an Addington (1971), dass größere Variabilität in der Grundfrequenz Sprecher überzeugender erscheinen lässt.

Research on pitch variation has shown that greater variety enhances competence, character, and sociability. (Burgoon et al. 1990:147)

Ähnliche Ergebnisse formulieren Scherer et al. (1973). In ihrer Studie beschreiben sie ein Experiment, in dem ein Text von einem Schauspieler in zwei verschiedenen Sprechfassungen gelesen wird: einmal in einer paralinguistisch stimmigen Version und einmal in einer paralinguistisch unstimmigen Version. Diese Texte werden Studenten vorgespielt, die die Sprechfassungen anhand verschiedener Skalen einschätzen sollen. Als Ergebnis lässt sich der folgende Befund festhalten:

Confident Text but not Confident Voice led to higher ratings on the adjectives ‚conceited‘, ‚professional‘, and ‚businesslike‘, and to lower ratings on the adjective ‚personal‘. [...] Confident Voice but not Confident Text led to higher ratings on the adjectives ‚enthusiastic‘, ‚forceful‘, ‚active‘ and ‚competent‘. It is interesting that competence is attributed on the basis of voice, while task-orientation [...] is attributed on the basis of linguistic cues, since these traits are usually seen as highly related. (Scherer et al. 1973:42)

Abschließend sollen zwei Untersuchungen von Gélinas-Chebat & Chebat (1992 und 1999) vorgestellt werden, die gleichzeitig eine Überleitung zum nächsten Abschnitt, der die Lautstärke zum Untersuchungsgegenstand hat, darstellen. Den theoretischen Rahmen dieser Studien bildet das ELM (s.o.), worin angenommen wird, dass periphere Signale wie prosodische Merkmale vor allem dann die Einstellung des Empfängers beeinflussen, wenn es sich um Situationen handelt, in die der Rezipient nicht so hoch involviert ist.

The ELM proposes that peripheral cues such as voice characteristics should enhance the receiver's attitude only under low-issue involvement situations. (Gélinas-Chebat & Chebat 1992:447)

Ist der Rezipient hingegen hoch motiviert, treten periphere Signale zu Gunsten einer Fokussierung verbalargumentativer Impulse in den Hintergrund. In beiden Studien untersuchen Gélinas-Chebat & Chebat den Einfluss von Grundfrequenz und Lautstärke in Bezug auf zwei Dimensionen der Glaubwürdigkeit innerhalb der Werbung, nämlich Verinnerlichung und Identifikation. Dabei finden sie heraus, dass Grundfrequenz und Lautstärke zu asymmetrischen Resultaten führen:

[...] intonation has a main effect on identification, which is not true for intensity. Conversely, the interactive effects of intensity and involvement were significant on both constructs, which was not true for intonation. (Gélinas-Chebat & Chebat 1999:334)

### 2.4.3 LAUTSTÄRKE

In Zusammenhang mit persuasiven Effekten der Lautstärke nennen Brown & Bradshaw (1985) große Lautstärke als Indikator für Sprechereigenschaften wie Dominanz, Selbstbewusstsein und Aggression. Dieses Phänomen beschreiben Page & Balloun bereits 1978. Sie ließen Interviewer mit verschiedenen Lautstärken von Hörern beurteilen und kamen zu dem Schluss:

The Interviewee was perceived as most aggressive when speaking in a high voice volume, but was also perceived to be lacking in self-assurance. (Page & Balloun 1978:65)

Einen anderen Aspekt der Lautstärke rücken Zuckermann & Driver (1989) und Zuckermann et al. (1990) in den Mittelpunkt ihrer Argumentation. Sie gehen eher davon aus, dass weniger die absolute Lautstärke als vielmehr die Lautstärkevariation innerhalb einer Äußerung die Wahrnehmung der sprecherischen Wirkung beeinflusst. Demnach können Sprecher mit variierender Lautstärke effektiver die Aufmerksamkeit auf sich ziehen und erscheinen somit attraktiver als Sprecher, die nur minimale Anzeichen der Lautstärkevariierung anzeigen. Beobachtungen wirkungsvoller Selbstdarsteller zeigen, dass Lautstärkevariation als ein Indiz für deren sprecherischen Erfolg angeführt werden kann.

[...] ‚salespeople‘ such as television evangelists, politicians, and well-known communicators including Paul Harvey, all of whom demonstrate considerable loudness variability in their vocal behaviour. (Peterson et al. 1995:4)

Eine Zusammenfassung verschiedener stimmlicher Parameter bietet die Studie von Sharf & Lehman (1984). In ihr wird das Zusammenwirken der akustischen Größen Sprechgeschwindigkeit, Grundfrequenz und Lautstärke mit dem Erfolg von Telefon-Interviewern und der Hörerperzeption in Zusammenhang betrachtet. Dabei werden sechs Interviews (drei Sprecher mit hoher Erfolgsrate und drei Sprecher mit niedriger Erfolgsrate) 18 Hörern vorgespielt, die diese im Anschluss einschätzen sollen. Folgende Ergebnisse halten die Autoren fest: (1) Erfolgreiche Interviewer haben höhere Grundfrequenzen und größere Variationen im Grundfrequenzbereich als weniger erfolgreiche Interviewer. (2) Die Sprechgeschwindigkeit (Silbe pro Sekunde) kann nicht als Kriterium für den Erfolg herangeführt werden, aber erfolgreichere Interviewer weisen kürzere Pausen auf. (3) Erfolgreichere Interviewer verwenden zum Eröffnen eines Ge-

sprächs terminale Intonationsmuster im Gegensatz zu erfolglosen Interviewern, die ihre Gespräche mit interrogativen Mustern eröffnen. (4) Signifikante Korrelationen bilden sich unter den verschiedenen Parametern der Sprechweise heraus, aber (5) es werden keine signifikanten Zusammenhänge zwischen dem Hörerurteil und der Erfolgsrate des Interviewers nachgewiesen.

## 2.5 ZUSAMMENFASSUNG

Die hier besprochenen Studien bieten einen Überblick über die Forschungslage zur Verbindung von Prosodie und Persuasion unter verschiedenen Gesichtspunkten. Alle Arbeiten kommen dabei zu dem Ergebnis, dass prosodische Marker als wichtiges Indiz für persuasive Effekte angeführt werden können, was für den gezielten Einsatz in Radio- und Fernsehwerbung von hoher Bedeutung ist. Aber es muss auch kritisch angemerkt werden, dass viele der Ergebnisse nicht ohne vorheriges Abwägen auf die aktuelle Forschung übertragen werden können. Einzelne Autoren verwenden (vor allem in älteren Arbeiten) sehr übersichtliche Datensätze, so dass die statistische Aussagekraft einiger Arbeiten durchaus kritisch gesehen werden muss. Auch birgt das Außerachtlassen mehrerer Parameter die Gefahr, dass das Zusammenwirken von stimmlichen und sprecherischen Parametern unberücksichtigt bleibt, welches mit Sicherheit Rückschlüsse auf persuasive Effekte gewährleistet. Stimmeigenschaften wie z.B. Timbre oder Sprechereigenschaften wie z.B. Akzentuierung bleiben in fast allen Untersuchungen unbeachtet. Letztlich ist auch die Frage danach, inwieweit es kulturelle Differenzen in der Produktion und in der Perzeption prosodischer Muster gibt, nicht dahingehend beantwortet, als dass alle Ergebnisse der besprochenen Studien unkritisch auf die hier vorliegende Arbeit übertragen werden könnten. Es werden vor diesem Hintergrund einige Fragen nach der Form und Funktion der prosodischen Mittel neu diskutiert werden müssen. Schließlich bleibt auch anzumerken, dass zentrale Aussagen und Formulierungen in einzelnen Aufsätzen eher vage und unpräzise bleiben, so dass einige Begriffe mangels fehlender Definition in Folgearbeiten nicht optimal aufgegriffen und präzise weiterverwendet werden können.

Als Fazit lässt sich formulieren, dass die referierten Arbeiten wichtige Tendenzen aufzeigen, die es durch begriffliche Präzisionen zu konkretisieren gilt. Aber

anbetracht noch ausstehender Forschungsarbeiten in der Schnittstelle phonetischer Untersuchungen von Persuasion und Prosodie erweisen sich die zitierten Studien als außerordentlich relevant für die hier vorliegende Frage danach, welche Bedeutung prosodischen Merkmalen im persuasiven Kontext der Werbung zukommt.

### 3 EMOTION UND PROSODIE

#### 3.1 EMOTIONS KONZEPTE

Die wissenschaftliche Auseinandersetzung mit Emotionen ist in den letzten Jahren in verschiedenen Disziplinen zu einem Forschungsgegenstand von erhöhtem Interesse geworden. Unter differenten Fragestellungen versuchen Linguisten, Psychologen, Neurowissenschaftler und auch Phonetiker und Sprechwissenschaftler zu beschreiben,

[...] was Emotionen sind, wie sie zustandekommen und welche Rolle sie im Gesamtzusammenhang des Erlebens und Verhaltens spielen. (Meyer et al. 1993:34)

Alle Teildisziplinen nähern sich dabei dem Phänomen „Emotion“ von verschiedenen Seiten, so dass es bisher keine konsensfähige Definition zu geben scheint und auch die Frage nach dem Sinn einer allgemeingültigen Definition gestellt werden muss (Meyer et al. 1993:22). Eine treffende Umschreibung dieser Definitionsproblematik benennt Schmidt-Atzert (1996:18) in Anlehnung an Wenger et al. (1962:3):

Emotion ist ein seltsames Wort. Fast jeder denkt, er versteht, was es bedeutet, bis er versucht, es zu definieren. Dann behauptet praktisch niemand mehr es zu verstehen.

Der Begriff „Emotion“ leitet sich vom lateinischen *emovere* ab, was soviel wie aufwühlen, herausbewegen, erschüttern oder her austreiben bedeutet. Er bezeichnet eine psychophysiologische Zustandsänderung, die durch äußere Reize (Sinnesempfindungen), innere Reize (Körperempfindungen) oder kognitive Prozesse (Bewertungen, Vorstellungen, Erwartungen) ausgelöst werden kann. Die daraus resultierenden emotionalen Reaktionen (emotional responses) gehen mit einer emotionalen Erregung (emotional arousal) einher. Sie sind gekennzeichnet durch den emotionalen Ausdruck (emotional expression) und die Desorganisation bzw. Unterbrechung des momentan ausgeführten Verhaltens (Fröhlich 2000:148). Inwieweit diese emotionalen Ausdruckskomponenten auch durch prosodische Markierungen transportiert werden, wird im Weiteren noch geklärt werden.

Bei Emotionen handelt es sich also um Prozesse, an denen mehrere psychische Komponenten beteiligt sind (Kranich 2003:41). Unstrittig ist auch, dass

sich Emotionen voneinander unterscheiden lassen, worin dabei aber Unterschiede und Gemeinsamkeiten liegen, lässt sich schwer aufzeigen (Otto et al. 2000:30). In ihren Ausführungen unterscheiden Otto et al. (2000:45f) allein zehn Ansätze zur Emotionsbeschreibung, die sich teilweise diametral gegenüber stehen. Einen Versuch, diese definitorische Vielfalt zu strukturieren haben Kleinginna & Kleinginna (1981) vorgenommen. Sie bildeten aus etwa 100 Definitionsversuchen einen Querschnitt und schlugen im Resultat folgende Definition als Konsens ihrer Untersuchung vor:

Emotion is a complex set of interactions among subjective and objective factors, mediated by neural/hormonal systems, which can (a) give rise to affective experiences such as feeling of arousal, pleasure/displeasure; (b) generate cognitive processes such as emotionally relevant perceptual effects, appraisals, labeling processes; (c) activate widespread physiological adjustments to the arousing conditions; and (d) lead the behavior that is often, but not always, expressive, goal-directed, and adaptive. (Kleinginna & Kleinginna 1981:355)

Definitorische Divergenzen ergeben sich auch aus unscharfen Differenzierungen zu benachbarten Termini wie Gefühl, Affekt und Stimmung. Nach Fröhlich (2000:148) werden die spürbar einsetzende Erlebnisweise und die Erlebnisqualität von Emotionen als Gefühl, und intensive, kurzzeitige Gefühle mit einengender Wirkung auf Erleben und Verhalten (z.B. Freudentaumel, Panikanfall) als Affekt bezeichnet. Längerandauernde Erlebnistönungen ohne klaren Reiz- oder Bedürfnisbezug (wie etwa Manie, Mutlosigkeit, Depression) werden als Stimmung bezeichnet. In älteren Quellen findet sich dafür auch die Verwendung des Begriffs Gemütsbewegung (Darwin 1901, Wundt 1911).

Zur Erklärung der Wirkungsweise von Emotionen wurden unterschiedliche Modelle entwickelt. Neben diskreten und dimensional Modellen wird in der aktuellen Literatur hauptsächlich Scherers Prozesskomponenten-Modell diskutiert (Kranich 2003:44f, Paeschke 2003:65f, Kienast 2002:12f, Kehrein 2002:105f).

Erstere, die Vertreter der diskreten Emotionsmodelle, versuchen konkrete Grund- oder Basisemotionen in ihren spezifischen Eigenschaften zu identifizieren und zu klassifizieren (Ekman 1992, Izard 1999, Schmidt-Atzert 1987, Plutchik 1980). Die Existenz solcher Basisemotionen begründen die Verfechter damit, dass diese Primäremotionen (a) ontogenetisch am frühesten auftreten, (b) dass sie konsistent (universell) in allen Kulturen wieder zu finden sind und dass sie (c) den größten evolutionären Anpassungswert für den Menschen (z.B. Furcht als Schutz vor Gefahr) haben (Mees 1991:170).

Ortony & Turner (1990) haben den Versuch unternommen, verschiedene Auffassungen in Bezug auf die Formulierung von Basisemotionen aufzulisten. Auch wenn an dieser Stelle nicht gezeigt werden kann

[...] about how many emotions are basic, which emotions are basic, and why they are basic (Ortony & Turner 1990:315)

so wird aber verdeutlicht, dass es zwischen den angenommenen Typologien sowohl quantitative als auch qualitative Differenzen hinsichtlich der zugrunde gelegten Emotionen gibt. Zudem scheint es schwierig, sich darüber einig zu werden, ob überhaupt alle aufgelisteten Emotionen als Gefühle bezeichnet werden können:

Nach der hier vorgelegten Analyse sind übrigens ‚Überraschung‘, ‚Interesse‘, ‚Erwartung‘, ‚Aufnahmebereitschaft‘, ‚Schüchternheit‘ *keine* Gefühle (Mees 1991:169)

Deutliche Übereinstimmungen in der Auflistung ergeben sich für die Emotionen Ärger und Wut, Furcht und Angst, Freude und Glück, Ekel und Trauer, welche von nahezu allen Autoren als Basisemotion angeführt werden. Homogenität herrscht auch in der Ansicht, dass durch Konditionierung aus Basisemotionen komplexere Emotionsgebilde entstehen können (Fröhlich 2000:148), wie etwa Eifersucht, Selbstmitleid oder Genugtuung. Dabei referiert Mees (1991:171f) zwei Modelle, die sich gegenüberstehen. Nach ersterem entsteht eine komplexe Emotion aus Grundemotionen wie in der Farbenlehre Mischfarben aus Grundfarben entstehen, wobei die zugrunde liegenden Emotionen beobachtbar bleiben. Nach der anderen Analogie werden die Komponenten von komplexen Emotionen mit der Zusammensetzung chemischer Elemente verglichen, wonach aus verschiedenen Elementen eine neue Verbindung besteht, deren Einzelteile nicht mehr separat beobachtbar sind.

Den diskreten Emotionsmodellen stehen dimensionale Erklärungsmodelle gegenüber. Diese Modelle gehen auf Wilhelm Wundt zurück, der bereits zu Beginn des 20. Jahrhunderts ein dreidimensionales System zur Charakterisierung von Gefühlen mit den Achsen (a) Lust/Unlust, (b) Erregung/ Beruhigung und (c) Spannung/Lösung vorschlägt (Wundt 1911:201, Scherer 1996:306). Diese drei Skalen lassen sich wiederum auf die Bewertungsdimensionen: Valenz (angenehm/unangenehm), Aktivität (erregend/ beruhigend) und Potenz (stark/schwach) übertragen (Grimm & Engelkamp 1981:19f). Eines der bedeutendsten Dimensionsmodelle entwickelte Schlosberg (1954), indem er die drei

Wundtschen Skalen aufgreift und ihnen die Polaritäten (a) Achtung – Abwehr, (b) positiv – negativ und (c) Schlaf – Spannung zuordnet. Der Kritikpunkt an den Dimensionsmodellen ergibt sich aus Zweifeln an der vorausgesetzten Unabhängigkeit der drei Emotionsdimensionen. Untersuchungen haben gezeigt, dass besonders die Valenzskala mit der Aktivitätsskala korreliert (Izard 1999:45).

Die am häufigsten angewendete Methode zur Überprüfung der Emotionsdimensionen ist die Klassifizierung von mimischen Ausdrucksformen, bei denen die Versuchspersonen Fotoserien mit verschiedenen Gesichtsausdrücken bewerten sollen (Ekman & Friesen 1971). Unabhängig vom Land, in dem oder der Sprache, in der die Untersuchungen stattfanden, bestätigten sich dabei immer wieder zwei von Wundts Skalen, die Valenz- und die Aktivitätsdimension (Scherer 1996:306f, Kienast 2002:13, Kehrein 2002:142).



**Abb. 4: Zweidimensionale Darstellung von Emotionsbezeichnungen: vertikale Dimension = Aktivität, horizontale Dimension = Valenz (Scherer 2003: 201.)**

Abbildung 4 gibt eine Variante dieser zweidimensionalen Darstellung von Emotionsbezeichnungen in Anlehnung an Scherer (1984:51) wieder. Die Ansichten zur Bedeutung der dritten Wundtschen Skala divergieren in der Literatur. Paeschke (2003:57f) nennt stellvertretend für die Potenzdimension (stark/schwach) folgende Alternativen: selbstbewusst/unsicher, dominant/unterwürfig oder Zuwendung/Abwendung.

Diese Form der zweidimensionalen Anordnung von Emotionsbezeichnungen (vgl. Abb. 4) bietet einen guten Überblick über Graduierungen und Übergänge von einzelnen Emotionskonzepten in andere, wenngleich an dieser Stelle nicht geklärt werden kann, inwieweit sich diese modellhafte Anordnung auch empirisch verifizieren lässt. Entscheidend für diese Arbeit, die sich mit emotionalisierten Sprechweisen in der Werbung beschäftigt, ist, dass sich aus der Fragestellung ein deutlicher Interessenfokus zur linken Hälfte, der die positiven Eigenschaften abbildet, mit den Dimensionen aktiv & passiv ausmachen lässt. Im Mittelpunkt steht dabei nicht die Frage nach graduellen Unterschieden, wie z.B., ob eine Äußerung als jubelnd, als begeistert oder überrascht eingeschätzt wird, sondern das Anliegen, im Werbekontext positive Grundaussagen und Emotionen zu unterbreiten und hervorzurufen. Inwieweit stimmliche und sprecherische Merkmale diese Färbungen auslösen können, wird an späterer Stelle genauer beschrieben werden.

Vielseitige Akzeptanz findet in der aktuellen Literatur zur Emotionsforschung Scherers Komponenten-Prozess-Modell (Scherer 1986:143ff, Zentner & Scherer 2000:158f, Kienast 2002:12ff, Kranich 2003:44ff, Kehrein 2002:105ff, Paeschke 2003:65f), wobei der Autor davon ausgeht, dass die Bewertung von Ereignissen in sequenziell angeordneten Reizverarbeitungsschritten verläuft, wonach jeder Organismus externe und interne Reize mittels so genannter Stimulus Evaluation Checks (SEC's) prüft. Den Ausgangspunkt für die Entwicklung dieses Modells bildete Scherers Schwierigkeit mit den diskreten Emotionsmodellen, den Output einer emotionalen Episode vorherzusagen. Nach seiner Modellvorstellung stellen Emotionen dynamisch aktive Prozesse dar, wobei die Prüfung und Bewertung von Reizen in unterschiedlichen Subsystemen stattfindet und von daher differenzierte Aussagen über Zustandsänderungen im Organismus zulässt.

[...] Scherer [...] has suggested a sequence theory of emotional differentiation, which postulates that each organism's information processing subsystem continuously scans external and internal stimulus input and performs a series of

stimulus evaluation checks (SECs) using functionally defined criteria. (Scherer 1986:147)

Nach Scherers Theorie erfolgt die Situations- und Objektbewertung immer anhand folgender fünf (stets in derselben Reihenfolge ablaufenden) SEC's (Zentner & Scherer 2000:158f):

<b>SEC</b>	<b>Komponenten</b>
<p><b>(1) Neuheit</b> (Novelty check)</p>	<ul style="list-style-type: none"> <li>▪ es wird festgestellt, ob Veränderungen bei inneren oder äußeren Reizen auftreten, insbesondere ob ein neues Ereignis eintritt oder zu erwarten ist</li> </ul>
<p><b>(2) Intrinsische Angenehmheit</b> (Intrinsic pleasantness check)</p>	<ul style="list-style-type: none"> <li>▪ es wird evaluiert, ob ein Stimulus angenehm ist und Annäherung nahe legt oder unangenehm ist und zu Vermeidung anregt</li> </ul>
<p><b>(3) Zieldienlichkeit</b> (Goal/need significance check)</p>	<ul style="list-style-type: none"> <li>▪ es wird geprüft, ob ein Stimulus für das Erreichen von Zielen relevant ist</li> <li>▪ und ob ein Resultat mit bestimmten Erwartungen in der geplanten Handlungssequenz übereinstimmt oder nicht</li> <li>▪ und ob ein Ereignis die Zielreichung fördert oder nicht</li> </ul>
<p><b>(4) Bewältigungsvermögen</b> (Coping potential check)</p>	<ul style="list-style-type: none"> <li>▪ es wird geprüft, wodurch das relevante Ereignis verursacht wurde</li> <li>▪ und ebenso das Bewältigungsvermögen, über das der Organismus verfügt</li> <li>▪ und weiterhin das relative Vermögen des Organismus, das Ereignis zu verändern (z.B. vermeiden oder neutralisieren)</li> <li>▪ und schließlich das Vermögen, sich an das Endresultat aufgrund interner Restrukturierungsprozesse anzupassen</li> </ul>
<p><b>(5) Selbst/Norm Kompatibilität</b> (Norm/self compatibility check)</p>	<ul style="list-style-type: none"> <li>▪ es wird evaluiert, inwieweit das Ereignis mit sozialen Normen, kulturellen Konventionen übereinstimmt</li> <li>▪ und ob Kompatibilität mit eigenen internalisierten Normen und Standards besteht oder nicht</li> </ul>

**Tab. 3: Scherers Komponenten-Prozess-Modell (SECs)**

Ein spezifischer emotionaler Zustand wird nach diesem Modell als Endresultat der einzelnen Komponenten im Prozess gesehen, beispielsweise

[...] ließe sich Furcht als emotionale Resultante von Bewertungen eines Ereignisses als neu, unangenehm, zielhindernd und nur durch Flucht zu bewältigen definieren. Wenn allerdings die Evaluation des Bewältigungsvermögens anzeigt, dass der Organismus über ausreichende Ressourcen zu einem Erfolg versprechenden Angriff verfügt, dann dürfte der Ärger die emotionale Reaktion auf dasselbe Ereignis sein. (Zentner & Scherer 2000:159)

Im Fokus dieser Arbeit steht die Frage danach, wie in einem konkreten Kontext (Werbung) die möglichen Resultate emotionaler Episoden auch prosodisch markiert sein können, also die Frage nach der sprecherischen und stimmlichen Ausdrucksgestaltung, was in der Literatur auch als vokaler Emotionsausdruck beschrieben wird.

### 3.2 VOKALER EMOTIONS-AUSDRUCK UND PROSODIE: DAS LINSENMODELL VON BRUNSWIK IN EINER ADAPTION VON SCHERER

Das Verständlichste an der Sprache ist nicht das Wort selber,  
sondern Ton, Stärke, Modulation, Tempo,  
mit denen eine Reihe von Wörtern gesprochen wird,  
kurz, die Musik hinter den Worten,  
die Leidenschaft hinter dieser Musik,  
die Person hinter dieser Leidenschaft:  
alles das also, was nicht geschrieben werden kann.

Friedrich Nietzsche<sup>7</sup>

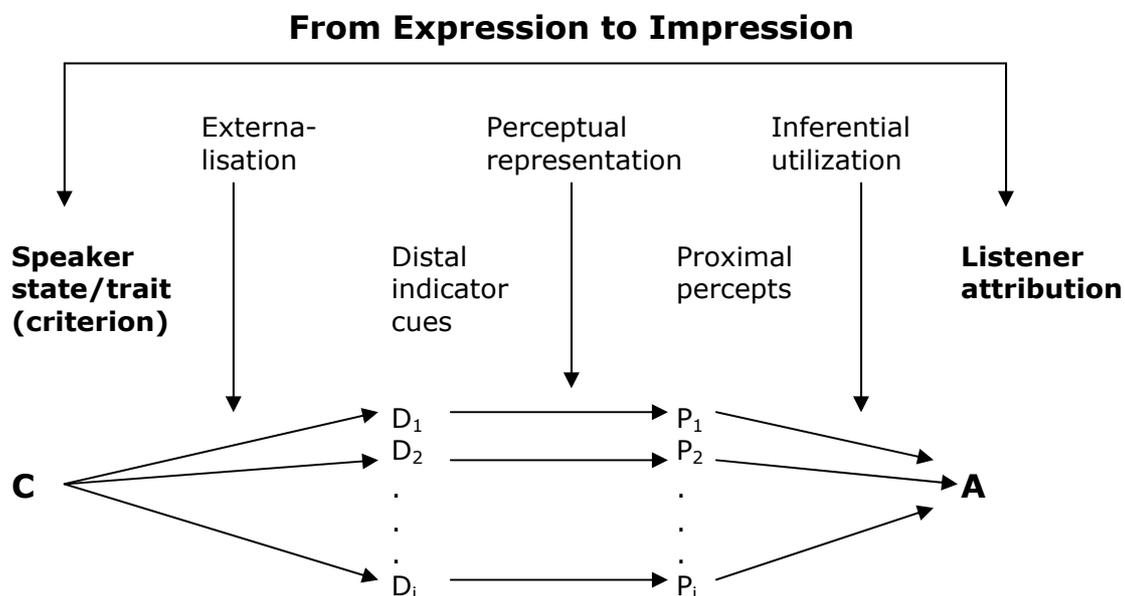
Die prosodischen Einheiten der Sprache leisten mehr als das bloße Überbringen semantischer Inhalte, sie drücken immer auch etwas über die momentane Sprecherbefindlichkeit aus. Stimme und Sprechweise sind somit ein Medium, um etwas über die Wesensart eines Sprechers zu kommunizieren. Das sprachliche Signal birgt dabei eine Fülle an Informationen über Alter, Geschlecht, regionale (und soziale) Herkunft und Stimmungslage eines Sprechers. Außerdem gibt es Auskunft über Beziehungsverhältnisse zwischen Personen wie Sympathie und Antipathie oder Dominanz und Unterwürfigkeit.

---

<sup>7</sup> aus: Nietzsche (1978:190f)

Kramer [...] reviewed a number of studies which have demonstrated that various aspects of a speaker's physical and emotional state, including age, sex, appearance, intelligence, and personality can be identified by voice alone. (Murray & Arnott 1993:1099)

In der Literatur finden sich zur Beschreibung prosodischer Ausdrucksphänomene zwei methodisch differente Herangehensweisen. Die eine Richtung verfolgt aus der Sprecherperspektive das Ziel, objektive Aussagen über prosodische Veränderlichkeiten in der Stimme und Sprechweise anhand der Beschreibung konkreter akustischer Parameter nachzuweisen. Auf der anderen Seite stehen Arbeiten, die aus der Hörerperspektive und durch Perzeptionstests versuchen, Aussagen über die Wahrnehmung emotionalisierter Sprechweisen zu treffen. Eine Synthese aus beiden Herangehensweisen ermöglicht das Linsenmodell von Brunswik (1956), welches Scherer (1978) in einer modifizierten Form für die vokale Charakterisierung des Emotionsausdrucks auf der einen Seite und die Wahrnehmung von Emotionen auf der anderen Seite adaptiert. Abbildung 5 zeigt eine vereinfachte Darstellung des Brunswikschen Linsenmodells.



**Abb. 5: Brunswiks Linsenmodell in einer Adaption von Scherer (1978:469, 2000a:2). Die Darstellung bezieht sich nur auf den Ausschnitt, den Scherer als „phenomenal level“ in sein Modell einführt. Aus Gründen der Prägnanz wird der Abschnitt „operational level“ ausgespart.**

Scherer (1978) greift dieses Modell auf, um einen methodischen Weg aufzuzeigen, der es erlaubt, Schlussfolgerungen über den Emotionszustand einer Person anhand stimmlicher und sprecherischer Merkmale abzuleiten. Diesen Inferenzprozess untergliedert er dazu in vier auseinander resultierende Bestandteile: (a) die Sprecherbefindlichkeit (speaker state/trait), (b) die objektiven Parameter, die durch die Sprecherbefindlichkeit beeinflusst werden (distale indicator cues), (c) die subjektiv wahrgenommenen Korrelate dieser Parameter (proximal percepts) und (d) die zugeschriebene Sprecherbefindlichkeit (listener attribution). Die Phasen, die sich als Bindeglieder zwischen den einzelnen Punkten ergeben (Externalisation, Perceptual representation & Inferential utilization), verdeutlichen die Prozesshaftigkeit des Modells.

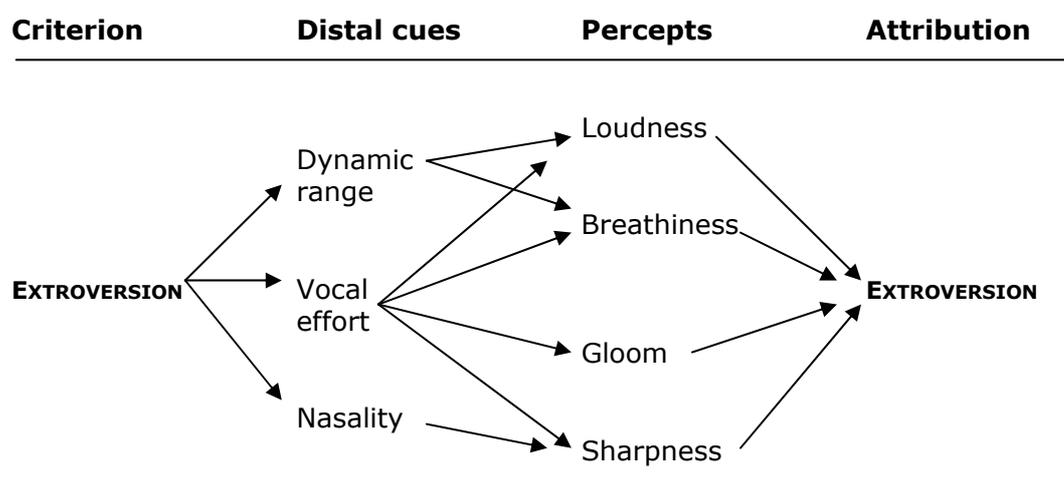
Abbildung 5 schematisiert zudem die Darstellung des Expressions- und Perzeptionsprozesses zweier Personen, wobei eine Person (A) die emotionalen Attribute einer anderen Person (C) anhand externer Sprechermerkmale inferriert. Dabei wird ein Sprecher (C) durch habituelle und aktuelle Merkmale näher spezifiziert. Bei den habituellen Merkmalen handelt es sich um Merkmale der Persönlichkeit oder Einstellung, die gewöhnlich als sehr stabil eingeschätzt werden, im Gegensatz zu den aktuellen Merkmalen, die vorrangig situationsabhängig und variabel sind, was im Besonderen auf den Ausdruck von Emotionen zutrifft. Beide Merkmalsbereiche können durch die Stimme und Sprechweise eines Sprechers ausgedrückt werden (Bergmann 1984:69). Die stimmlichen und sprecherischen Merkmale, die sich durch eine phonetisch-akustische Analyse objektiv erfassen lassen, werden als distale Hinweisreize bezeichnet. Im Modell wird dieser Abschnitt als Äußerlichkeit (Expression) bzw. Externalisation bezeichnet. Mit der Feststellung einzelner Parameter ist aber noch nichts darüber gesagt, ob diese auch Relevanz für den Hörer besitzen. Es ist deswegen (a) danach zu fragen, ob diese Hinweisreize, die im Modell (Abb. 5) als proximale Hinweisreize bezeichnet werden, vom Hörer wahrgenommen werden. Dieses Bindeglied bezeichnet Scherer als wahrgenommene Repräsentation (perceptual representation). Und es ist (b) danach zu fragen, ob die wahrgenommenen Reize vom Hörer auch auf die Sprecherbefindlichkeit attribuiert werden können. Diese Phase ist durch die gefolgerte Nutzbarmachung (inferential utilization) der Hinweisreize charakterisiert.

Scherer (1978) wählt die Synthese aus objektiven und subjektiven Untersuchungsmethoden (akustische Analyse und Perzeptionstest), um der Frage

nachzugehen, ob sich aus der Stimmqualität eines Sprechers Rückschlüsse auf die Persönlichkeit eines Sprechers ziehen lassen. Jede Kategorie wird dabei unabhängig von den anderen untersucht. Korrelationen zwischen den angrenzenden Kategorien korrespondieren mit dem Effekt, den eine Kategorie auf die nächste ausübt. Diese Annahmen stützt Scherer mit empirischen Untersuchungen. Am eindrücklichsten belegt er dies am Beispiel der Extrovertiertheit:

For the inference of extroversion [...] the following conditions are met: (a) the criterion is associated with ecologically valid voice energy cues (vocal effort and dynamic range), (b) these indicator cues are adequately represented as proximal voice percepts (particularly loudness and sharpness), and (c) percept utilization in the judges' inferential strategy corresponds to the association between criterion and distal indicator cues. (Scherer 1978:467)

In Abbildung 6 soll am Beispiel der Extrovertiertheit aufgezeigt werden, wie das Brunswiksche Modell in den Forschungskontext zur Untersuchung von vokal Hinweisreizen übertragen werden kann. Als Probanden wurden 24 männliche Amerikaner herangezogen, die in simulierten Jury-Diskussionen als Stimuluspersonen dienten. Die distalen Hinweisreize wurden durch 6 Phonetiker, die proximalen Hinweisreize durch 10 naive Beurteiler ermittelt. Die Höreinschätzung gegenüber der Sprecherbefindlichkeit erfolgte ebenfalls durch naive Hörer.



**Abb. 6: Inferenzprozess am Beispiel von Extrovertiertheit (Scherer 1978: 478). Auf die genaue Angabe der Korrelationen wird in dieser Übersicht verzichtet.**

Neben der Darstellung von Inferenzprozessen anhand von Stimmqualitätsmerkmalen bildet das Brunswiksche Linsenmodell aber auch einen theoretischen

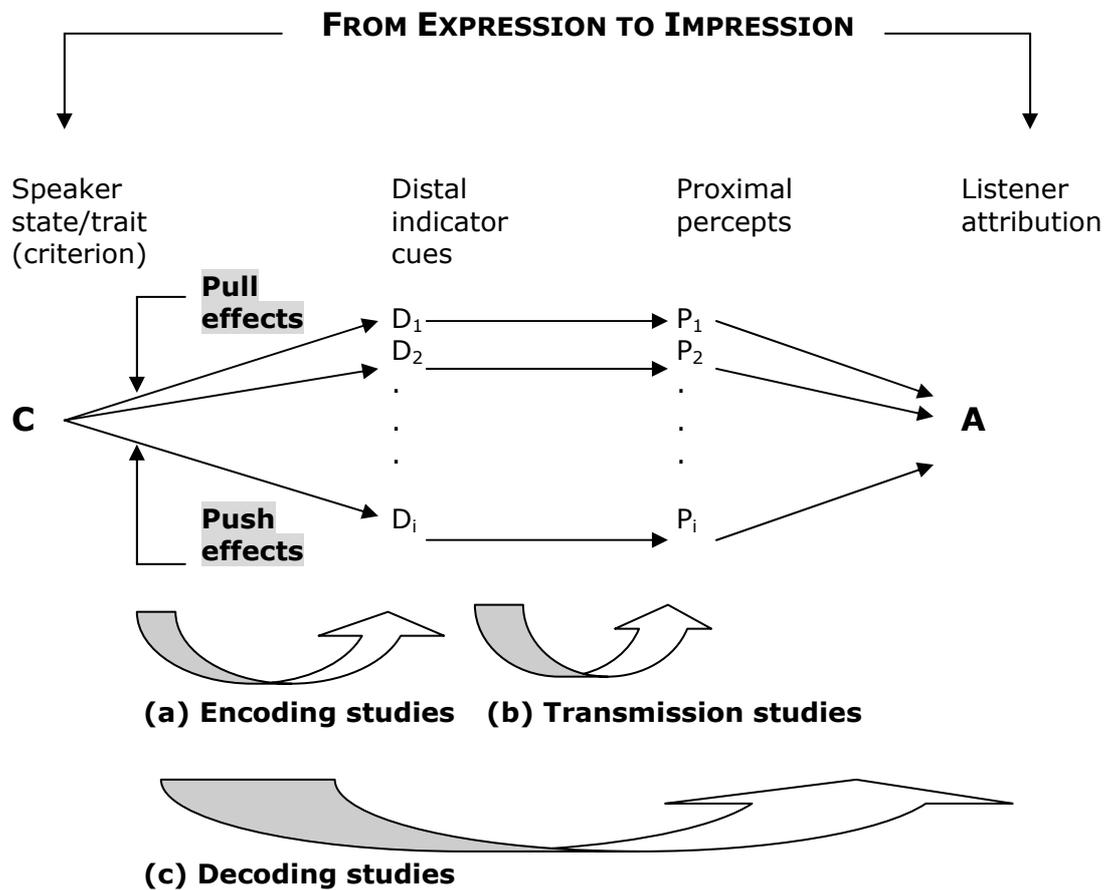
schen Bezugsrahmen, um mögliche Forschungsansätze und methodische Vorgehensweisen zur Beschreibung von vokalen Ausdrucksphänomenen voneinander abzugrenzen. Scherer (2000a:2) weist seiner adaptierten Schematisierung des Brunswikschen Modells (Abb. 5) die Überschrift „from expression to impression“ zu und steckt durch diese Begrifflichkeiten die Eckpunkte ab, zwischen denen sich die Forschungsliteratur einordnen lässt. Arbeiten im Schnittbereich von Stimme und Sprechweise und Emotion beschäftigen sich immer mit der Kombination aus Enkodierungs- und Dekodierungsaspekten. Dies geschieht, indem akustische Korrelate gemessen und unabhängig davon durch Perzeptionstests überprüft werden, so dass objektiv und subjektiv gewonnene Datensätze im Ergebnis synthetisiert werden können. Die meisten Studien fokussieren in ihren Arbeiten allerdings einen Aspekt stärker.

### 3.3 EMOTION UND PROSODIE: FORSCHUNGSÜBERBLICK

Die Schnittstelle von Emotion und Prosodie ist in den letzten Jahren wiederholt Fokus verschiedener Arbeiten gewesen, und auch aktuelle Studien stellen sich immer wieder Fragen nach dem Zusammenwirken von vokalem Ausdrucksverhalten und dem Identifizieren bestimmter Emotionen.

Die folgende Spezifizierung des Brunswikschen Modells in Abbildung 7 soll eine Systematik für die Einordnung der Forschungsliteratur darstellen. Danach lassen sich vor allem drei wissenschaftliche Herangehensweisen zur vokalen Ausdrucksgestaltung von Emotionalität unterscheiden, wobei in den meisten Arbeiten Elemente verschiedener Methoden zusammengeführt werden: (a) Arbeiten, die den Enkodierungsvorgang in den Mittelpunkt ihres Interesses rücken, (b) Arbeiten, die sich im Schnittpunkt von Enkodierung und Dekodierung befinden und damit den Übergang von prosodischen Parametern zu Perzeptionskonzepten beschreiben und (c) Arbeiten, die gezielt das Dekodieren von Emotionen durch bestimmte prosodische Markierungen fokussieren. Diese Einteilung nimmt auch Schröder (2004:37) in einer Vereinfachung vor, wenn er von sprecherzentrierten Studien (Enkodierung) und hörerzentrierten Studien (Dekodierung) spricht.

Vor allem jüngere Arbeiten synthetisieren oftmals unterschiedliche Herangehensweisen und untersuchen sowohl Enkodierungs- als auch Dekodierungsaspekte.



**Abb. 7: Das Brunswiksche Modell in einer Adaption Scherers zur Systematisierung der Forschungslage (Scherer 2000a:6f).**

Den Prozess der Verbindung der Herangehensweisen von Enkodierung über Transmission zu Dekodierung beschreibt Scherer (2000a:2f) mit dem Ausdruck „from expression to impression“ und bietet damit Kategorien an, die eine gute Systematisierung von wichtigen Überblicksarbeiten und Studien erlauben. Dieses System wird für die Aufarbeitung der Forschungsliteratur in den folgenden Kapiteln zugrunde gelegt.

### 3.3.1 ‚PUSH‘ UND ‚PULL‘ FAKTOREN

Das Ziel expressionistischer Arbeiten (z.B. Trojan 1952) ist es, sprecherische und stimmliche Parameter zu ermitteln und zu messen, die Rückschlüsse auf den emotionalen Zustand einer Person erlauben. Dabei wird das Enkodieren von Emotionen durch zwei Faktoren beeinflusst, die Scherer & Wallbott (1990:353ff) als Push- bzw. Pull-Faktoren bezeichnen.

**Push-Faktoren** beschreiben Veränderungen im organischen System, die direkte Auswirkungen auf die Art und die Intensität des motorischen Ausdrucksverhaltens haben. Es handelt sich dabei um psychobiologische Mechanismen, die auf physiologische Emotionsaspekte des zentralen Nervensystems (ZNS) und auf das vegetative Nervensystem zurückgehen. Eine sympathische Erregung wie z.B. bei Angst, Ekel oder Ärger ist darauf ausgerichtet, den Körper schnell zum Handeln zu bringen. Dies geschieht z.B. durch die Erhöhung der Herzschlagrate und des Blutdrucks, durch Veränderungen des Atemrhythmus oder durch die Verringerung des Speichelflusses. Eine parasympathische Erregung hingegen, wie sie etwa bei Trauer oder Wohlbefinden stattfindet, zieht durch einen eher verlangsamten Herzrhythmus, verstärkte Durchblutung der Eingeweide und eine erhöhte Speichelproduktion einen Zustand von Ruhe oder Entspannung nach sich. Auch könnte

[...] eine durch Hormonsekretion hervorgerufene Änderung des Speichelflusses durch Beeinflussung der Resonanzeigenschaften des Mundraumes die Stimmqualität verändern. (Scherer & Wallbott 1990:354)

Es gibt bisher keine Untersuchungen zum direkten Zusammenhang von vegetativer Erregung auf die Stimm- und Sprachproduktion (Kienast 2002:42), aber es ist anzunehmen, dass es einen indirekten Einfluss gibt, der sich beispielsweise über die Atmung auf die Sprachproduktion auswirkt.

Eine sympathische bzw. parasympathische Erregung führt zu einer Veränderung des Muskeltonus, das betrifft auch die an der Stimm- und Sprachgebung beteiligten Muskeln. Durch verschiedene Spannungsverhältnisse der an der Stimmgebung beteiligten Muskeln lassen sich unterschiedliche Phonationstypen diskriminieren, die Laver (1980:109f) auf folgende drei Parameter reduziert:

- |     |                              |  |
|-----|------------------------------|--|
| (a) | adduktive Spannung<br>(AS)   | = Spannung der Interarytenoid-Muskeln, die bei Kontraktion die Stellknorpel zusammenziehen.        |
| (b) | mediale Kompression<br>(MK)  | = Kontraktion der lateralen Cricoarytenoid-Muskeln, die zur Annäherung der Vocalisfortsätze führt. |
| (c) | Longitudinalspannung<br>(LS) | = Kontraktion der Vocalis-Muskeln und der Cricothyroid-Muskeln.                                    |

Kienast (2002:43) vermutet, dass die Folge von Emotionen mit sympathischer Erregung (Angst, Ekel, Ärger) eine stärkere Anspannung und Beweglichkeit auch der an der Stimm- und Sprachbeteiligung mitwirkenden Muskeln sein könnte. Bei extremer Anspannung kann dies bis zur Engebildung im laryngealen und supralaryngealen Bereich führen. Als Beispiel für eine hohe adduktive Spannung und eine hohe mediale Kompression kann die Knarrstimme angeführt werden (Eckert & Laver 1994:68), die von Ladefoged (1971:14f) wie folgt charakterisiert wird:

Another mode of vibration of the vocal cords occurs in laryngealized sounds. In this type of phonation the arytenoid cartilages are pressed inward so that the posterior portions of the vocal cords are held together and only the anterior (ligamental) portions are able to vibrate. The result is often a harsh sound with a comparatively low pitch. It is also known as vocal fry and creaky voice.

Im Gegensatz dazu sind Emotionen mit parasympathischer Erregung (Trauer, Wohlbefinden) durch eine geringere Muskelaktivität gekennzeichnet, die je nach Anspannungsgrad bis hin zur totalen Erschlaffung oder Lähmung führen kann. Symptomatisch für eine Entspannung der Kehlkopfmuskulatur ist die behauchte Stimme. Bei sehr geringer medialer Kompression und sehr geringer adduktiver Spannung schließt sich die Glottis nie ganz, so dass durch das Entweichen „wilder Luft“ der Luftverbrauch sehr ineffizient ist (Eckert & Laver 1994:76f).

Die **Pull-Faktoren** bezeichnen im Gegensatz zu den Push-Faktoren keine Begleiterscheinungen organischer Veränderungen, sondern vielmehr soziokulturelle Konventionen, wobei das Verhalten auf ein spezifisches Zielmuster gerichtet wird. Es handelt sich dabei um externe Vorgaben, die sich durch soziale Normierungen herausgebildet haben, wie beispielsweise die Intention, jemandem gefallen zu wollen. Dies kann auch für prosodische Markierungen gelten:

Pull effects [...] are externally based, in the sense that an externally defined vocal target is aimed at, be it for optimal signal transmission (a voice which 'carries well') or for social communication. (Schröder 2004:74)

Zu den soziokulturellen Konventionen zählen des Weiteren Faktoren, die dazu dienen, den Emotionsausdruck zu unterdrücken oder andere Emotionen vorzutäuschen, wie z.B. beim sozialen Lächeln. Diesen Kontrollaspekt, der willentlich gesteuert ist, hat Wundt bereits wie folgt beschrieben:

Indem der Culturmensch den Ausdruck seiner Affecte nach den Mitmenschen richtet, von denen er sich beobachtet weiß, sucht er mehr und mehr auch Gebarden und Mienen dieser Rücksicht anzupassen. Er sucht gewisse Affecte zu verbergen und andere auszudrücken. (Wundt 1905:85)

In der kulturgeschichtlichen Entwicklung haben sich Unterschiede in der Bedeutung akustischer Merkmale herausgebildet. Die Beschäftigung mit den paralinguistischen Attributierungen solcher Merkmale, zählt ebenfalls zu den Inhalten der Pull-Faktoren. So wird beispielsweise die Nasalität in unterschiedlichen Kulturkreisen verschieden interpretiert, was Laver (1980) anhand folgender Beispiele ausführt:

[...] in Cayuvava, a language of Bolivia, nasality is used stylistically with an honorific function: 'an individual of lower social or economic status addresses one of higher rank with a prominence of nasalization for all vowels of the utterance; and similarly with a women being polite to her husband [...]' (Crystal 1970:191). Finally, nasality is a very common setting component of voice quality, either idiosyncratically or as an indicator of membership of particular sociolinguistic groups. (Laver 1980:4)

Im Kontrast dazu, verfolgt Ohala (1983, 1984, 1994) in diversen Arbeiten einen ätiologischen und kulturübergreifenden Ansatz, indem er die Existenz eines so genannten „frequency codes“ annimmt, wonach scheinbar ungleiche Phänomene eine tiefer liegende Beziehung eingehen. Diese Überlegungen macht er an Parametern wie z.B.  $F_0$  fest, denen sich seiner Meinung nach laut-symbolische Entsprechungen zuordnen lassen. So werden tiefe  $F_0$  Werte mit körperlich großen Individuen und raue Stimmen mit Eigenschaften wie Stärke, Größe und Selbstzufriedenheit verbunden.

The sound-meaning correlations found in these cases adhere to the 'frequency code,' which also governs the vocalisations of other species, namely, where high  $F_0$  signifies (broadly) smallness, non-threatening attitude, desire for the goodwill of the receiver, etc., and low  $F_0$  conveys largeness, threat, self-confidence and self-sufficiency. (Ohala 1994:343)

Die Pull-Faktoren stellen für diese Arbeit einen wichtigen Bezugsrahmen dar, weil es bei der Frage nach der Bedeutung von vokalen Ausdrucksphänomenen

in der Werbung an erster Stelle um die Darstellung soziokulturell erlernter und willentlich eingesetzter Emotionen geht, die in der Regel in ihrer Realisierung und Perzeption von authentischen Emotionen, wie sie etwa bei den Push-Faktoren beschrieben werden, stark abweichen. In der Literatur wurde der Zusammenhang von stimmlichen Ausdrucksphänomenen in der Werbung bisher nicht näher untersucht, weswegen im Folgenden Studien vorgestellt werden, die sich dieser Problematik aus anderen Kontexten heraus zugewendet haben und deren Methoden und Ergebnisse sich auf die eigene Untersuchung übertragen lassen.

### 3.3.2 SPRECHERORIENTIERTE ARBEITEN: ‚ENCODING STUDIES‘ UND ‚TRANSMISSION STUDIES‘

In diesem Kapitel sollen ausgewählte Arbeiten aufgeführt werden, die sich aus der Sprecherperspektive mit dem Encodieren von Emotionen durch stimmliche Ausdrucksphänomene beschäftigen. Dies geschieht methodisch, indem die Sprecherbefindlichkeit (speaker state) kontrolliert wird, wobei Variablen der Sprechweise gemessen und mit der Sprecherbefindlichkeit in Verbindung gesetzt werden. Dabei handelt es sich um sehr unterschiedliche Ansätze wie die angeführten Beispiele zeigen werden. In den meisten Studien werden die emotionalen Sprechfassungen von Schauspielern (re-)produziert und mit einer neutralen Fassung kontrastiert.

In einer ersten Arbeit von Fernandez & Picard (2000) untersuchen die Autoren Veränderungen der Sprechweise in Stresssituationen. Solche Situationen erzeugen sie in einem Experiment, indem sie ihre Probanden, während diese in einem Fahrsimulator unterwegs sind, Matheaufgaben rechnen lassen. Der Stresslevel wurde optional durch die Fahrgeschwindigkeit (schneller 120 m.p.h./ langsamer 60 m.p.h.) oder durch die Frequenz der Matheaufgaben (alle 9 Sekunden oder alle 4 Sekunden) variiert. Die Sprechweisen von 4 Probanden wurden aufgezeichnet und in 4 Untergruppen analysiert (FF, SF, FS, SS). Dabei steht der erste Buchstabe jeweils für die Fahrgeschwindigkeit (fast/slow) und der zweite Buchstabe für die Frequenz der Rechenaufgaben (fast/slow). Für die Untersuchung der Sprechweisen wurden spektrale Frequenzbänder analysiert und automatische Klassifikatoren trainiert, um auf die-

se Weise Vorhersagen über den Grad von Stresszuständen anhand der Sprechweise zu ermöglichen.

In einer anderen Untersuchung von Amir et al. (2000) entwerfen die Autoren einen Kriterienkatalog zur automatischen Erkennung von fünf Basisemotionen (Ärger, Angst, Trauer, Freude und Ekel) anhand von drei physiologischen Variablen (Electromyogramm des Muskulus Corrugator (bei dessen Kontraktion ziehen sich die Augenbrauen zusammen), Herzschlagrate und sympathischer Hautreaktionstest), die gemessen werden, während die Probanden jeweils ein Ereignis für die entsprechenden Emotionen in ihre Erinnerung zurückholen und darüber berichten sollten. Die Sprechdaten wurden aufgezeichnet und analysiert. Dabei basiert die Methode auf der Extraktion von Grundfrequenzverläufen und Intensitätsstatistiken in bestimmten Fensterausschnitten. Als Ergebnis halten die Autoren fest:

Our results show that discrimination based on pitch average alone is much poorer than that obtained using the entire feature vector. Excluding, however, pitch average information decreases discrimination to a great extent. (Amir et al. 2000:4)

Den Zusammenhang von Sprache und Emotion und deren Identifikation anhand akustischer Korrelate untersuchen in den 70er Jahren bereits Williams & Stevens (1972), indem sie Schauspieler Dialoge kurzer Szenarien sprechen lassen, die verschiedene emotionale Stimmungen hervorrufen. Ausgewählte Ausschnitte der Aufnahmen wurden sowohl quantitativen als auch qualitativen Analysen unterzogen. Des Weiteren wurden die simulierten Sprachaufnahmen mit Spontanmitschnitten ähnlicher emotionaler Ereignisse aus dem realen Leben in Beziehung gesetzt und hinsichtlich einzelner Parameter verglichen. Die Daten wurden für die konkreten Emotionen Ärger, Angst und Trauer im Kontrast zu einer neutralen Variante in Bezug auf unterschiedliche akustische Parameter beschrieben. Zusammenfassend formulieren die Autoren:

Anger, fear, and sorrow situations tended to produce characteristic differences in contour of fundamental frequency, average speech spectrum, temporal characteristics, precisions of articulation, and waveform regularity of successive glottal pulses. Attributes for a given emotional situation were not always consistent from one speaker to another. (Williams & Stevens 1972:1238)

In einer ähnlich angelegten Arbeit untersuchen Tolkmitt & Scherer (1986) vokale Parameter im Zusammenhang mit induziertem Stress, der dabei sowohl kognitiven als auch emotionalen Ursprungs sein kann. Die Auswirkungen des Stresses werden in Beziehung zu Veränderungen der Phonation und der Arti-

kulationsprozesse während der Sprachproduktion gesetzt und separat für männliche und weibliche Sprecher analysiert, unter der Berücksichtigung von drei differierenden Persönlichkeitsprofilen: Sprecher mit niedriger Ängstlichkeit, Sprecher mit hoher Ängstlichkeit und Sprecher, die Ängstlichkeit ablehnen oder bestreiten. Anhand digitaler Aufnahmen werden die mittlere Grundfrequenz und Grundfrequenzverläufe, die Lage der Formanten und die spektrale Energieverteilung näher untersucht. Als Resultat formulieren sie:

Although mean  $F_0$  did not covary with stress manipulation,  $F_0$  floor of high-anxious and anxious-denying subjects increased with stress, probably due to physically based changes in muscle tension. For articulatory processes, as measured through formant location and spectral composition, significant changes were found for females. For anxiety-denying female subjects, precision of articulation increased under cognitive stress and decreased under emotional stress. (Tolkmitt & Scherer 1986:302)

Den Aspekt der Artikulationsgenauigkeit in Zusammenhang mit emotionalem Sprechausdruck untersuchen Kienast & Sendlmeier (2000) und Kienast (2002) für verschiedene Basisemotionen, wobei sie die Artikulation als in enger Beziehung zur Sprechgeschwindigkeit und zur Muskelspannung stehend, beschreiben. Emotionen mit geringeren Erregungsgraden (wie Trauer oder Langeweile) zeichnen sich dabei durch ein herabgesetztes Sprechtempo aus. Außerdem gilt, je höher die Sprechgeschwindigkeit, desto mehr ist eine Tendenz zur Vokalzentralisation zu beobachten. In ihren Untersuchungen versucht Kienast (2002:138) einen Zusammenhang zwischen der Artikulationsgenauigkeit einerseits und den Erregungsdimensionen andererseits (s.o.) herzustellen, wobei sie zu dem Ergebnis kommt, dass Emotionen mit geringerer Potenz wie Angst, Trauer und Langeweile auch mit einer geringeren Artikulationsgenauigkeit einhergehen, wohingegen z.B. Ärger, der als dominante Emotion auf der Potenzachse eingeordnet wird, auch eine wesentlich höhere Artikulationsgenauigkeit nach sich zieht.

Die Korrelation akustischer Parameter mit den drei Emotionsdimensionen testen auch Schröder et al. (2001) in ihrer Untersuchung. Als Ergebnis erhalten die Autoren dabei, dass nahezu alle akustischen Parameter substantielle Korrelationen mit den drei Emotionsdimensionen aufweisen. Am stärksten trifft dies für die Aktivitätsdimension zu:

The most numerous and strongest correlations were found for the activation dimension. Most acoustic variables correlate with activation, in the sense that expression of active emotion is accompanied by higher  $F_0$  mean and range,

longer phrases, shorter pauses, larger and faster F<sub>0</sub> rises and falls, increased intensity, and a flatter spectral slope. (Schröder et al. 2001:88)

Weniger stark ausgeprägt lassen sich aber auch Zusammenhänge zwischen akustischen Korrelaten zu den Dimensionen Potenz und Valenz beschreiben. Diese Verbindung stellt Davitz bereits 1964 her und schlussfolgert aus seinen Ergebnissen, dass Hörer Emotionen mit ähnlich starker Erregung wie Ärger und Freude leichter verwechseln als Emotionen mit unterschiedlichen Erregungsgraden. Pereira (2000) versucht diese Ergebnisse zu replizieren und kommt zu dem Schluss:

[...] that emotions with a similar level of arousal, and sometimes a similar level of power, share acoustic characteristics in terms of F<sub>0</sub> range and mean, and particularly intensity mean, and this must contribute to their perceived similarity and consequent confusion, especially in the hearing-impaired. (Pereira 2000:3)

Die meisten Arbeiten, die sich mit akustischen Variablen zur Kennzeichnung von Emotionen beschäftigen, bilden ihren Fokus auf den Dimensionen Aktivität und Erregung. Die Valenzdimension bleibt weitgehend unbeachtet, was als Problem mit sich bringt, dass sich beispielsweise Emotionen wie Freude und Ärger, die sich hauptsächlich hinsichtlich ihrer Valenz unterscheiden, schlecht diskriminieren lassen (Paeschke 2003:77). Da eine differenzierte Betrachtung einzelner Emotionen nicht Gegenstand dieser Untersuchung ist, wird im Folgenden darauf nicht genauer eingegangen.

In den Untersuchungen, die Paeschke & Sendlmeier (2000) und Paeschke (2003) zu prosodischen Charakteristika emotionaler Sprechweisen durchführen, erweitern sie den Katalog der untersuchten Parameter im Vergleich zu vorherigen Arbeiten um Merkmale wie Betonung und Rhythmus sowie das zur Charakterisierung des Stimmklangs gemessene Einschwingverhalten von Vokalen, die sich in Bezug auf die Diskriminierung einzelner Emotionen durchaus als aussagekräftig erweisen. In ihren Ergebnissen schließt sich Paeschke allerdings früheren Arbeiten an, ohne grundlegende Neuerungen zu beschreiben. Ähnlich wie bereits Williams & Stevens (1972, s.o.) untersuchen Abadjieva et al. (1993) akustische Merkmale des Emotionsausdrucks an einer Datenbank, die aus 18 Sprachaufnahmen zweier Schauspieler besteht. Auf der Grundlage von Oszillogrammen, Spektrogrammen, Frequenz- und Energiekonturen ermitteln die Autoren neben prosodischen Parametern wie Grundfrequenzverlauf und Sprechgeschwindigkeit auch Dauerwerte (Länge von betonten Silben) und

stimmqualitative Merkmale unter Berücksichtigung der Emotionen Ärger, Freude, Trauer, Angst und Ekel. Aufbauend auf diese Analyseergebnisse entwickeln die Autoren eine Software (HAMLET), die es ermöglicht, regelbasiert den Ausdruck des Text-To-Speech Synthesizers „DecTalk“ um einen emotionalen Ausdruck zu komplettieren. Die Ergebnisse des Syntheseverfahrens wurden im Anschluss in Hörtests überprüft. Weitere Arbeiten, die sich mit Sprachsyntheseverfahren beschäftigen, werden im Zusammenhang mit den „Decoding studies“ vorgestellt.

Abschließend sei noch auf zwei Überblicksarbeiten verwiesen, die aus einer ähnlichen Fragestellung heraus eine Auflistung von Studien, die nach prosodischen Parametern zur Beschreibung bestimmter Emotionen suchen, geben. Murray & Arnott (1993) stellen einen solchen Forschungsüberblick vor dem Hintergrund der Simulation von Emotionen in synthetisierter Sprache zusammen. Dafür erstellen die Autoren eine Tabelle (Tab. 4), in der stimmliche Effekte konkreter Emotionen im Vergleich zu einer relativen neutralen Sprechweise zusammengefasst sind:

	<b>Anger</b>	<b>Happiness</b>	<b>Sadness</b>	<b>Fear</b>	<b>Disgust</b>
<b>Speech rate</b>	slightly faster	faster or slower	slightly slower	much faster	very much slower
<b>Pitch average</b>	very much higher	much higher	slightly lower	very much higher	very much lower
<b>Pitch range</b>	much wider	much wider	slightly narrower	much wider	slightly wider
<b>Intensity</b>	higher	higher	lower	normal	lower
<b>Voice quality</b>	breathy, chest tone	breathy, blaring	resonant	irregular voicing	grumbled, chest tone
<b>Pitch changes</b>	abrupt, on stressed syllables	smooth, upward inflections	downward inflections	normal	wide, downward terminal inflections
<b>Articulation</b>	tense	normal	slurring	precise	normal

**Tab. 4: Zusammenfassung stimmlicher Emotionseffekte im Vergleich zu neutraler Sprechweise. (Murray & Arnott 1993: 1106).**

Frick (1985) stellt in seinen Ausführungen eine Literaturübersicht vor, die primär der Frage nachgeht, welche Rolle prosodischen Merkmalen im Kommunizieren von Emotionen zukommt. Dabei diskutiert er Korrelationen zwischen Emotionsdimensionen und prosodischen Merkmalen vor dem Hintergrund, dass bestimmte prosodische Konturen (vor allem Grundfrequenz- und Intensitätsverläufe) benutzt werden, um spezielle Emotionen auszudrücken. In einer späteren Arbeit (Frick 1986) konkretisiert er dies für den prosodischen Ausdruck am Beispiel von Ärger. Dabei nimmt er an, dass hohe Grundfrequenzen auf eher unterwürfige Sprecher hinweisen oder mit Freude korrelieren, wohingegen tiefere Frequenzlagen eher als aggressiv oder dominant empfunden werden.

Eine aktuelle Arbeit aus der kognitiven Neurologie beschäftigt sich mit der Frage nach dem Hören von Gefühlen und stellt somit ein Bindeglied zwischen den sprecherzentrierten und den hörerzentrierten Studien dar. Durch funktionell-bildgebende Verfahren versuchen Ackermann et al. (2004) und Schirmer & Kotz (2006) neuroanatomische Grundlagen zu beschreiben, die für die Verarbeitung affektiver Prosodie verantwortlich sind. Dabei weisen sie auf folgende zwei Stufen der Wahrnehmung affektiver prosodischer Einheiten hin:

[...] a) eine vorwiegend rechtshemisphärische Enkodierung von Intonationskonturen, ein akustisches Korrelat affektiver Prosodie, im Bereich posteriorer Anteile des Gyrus temporalis superior und b) eine im Wesentlichen an orbitofrontale Strukturen beidseits geknüpfte Evaluation von Art und Ausprägung der durch affektiv-prosodische Merkmale des Sprachsignals vermittelten Emotionen. (Ackermann et al. 2004:449)

Es sollen nun Studien vorgestellt werden, die sich mit dem Dekodieren von Emotionen beschäftigen und somit den methodischen Schwerpunkt auf die Hörerperspektive legen, wobei die Attribution von Sprecherzuständen durch Hörerurteile den Interessenfokus bildet.

### 3.3.3 HÖRERZENTRIERTE ARBEITEN: ‚DECODING STUDIES‘

Unter den „Decoding studies“ werden Arbeiten subsumiert, die sich inhaltlich mit dem Identifizieren von Emotionen anhand von Hörerurteilen beschäftigen. Damit unterscheiden sich diese Arbeiten methodisch von den bisher genannten insofern, als nicht simulierte Sprechfassungen verschiedener Emotionalitäten

untersucht, sondern Parameter manipuliert und die sich auf diese Weise unterscheidenden Fassungen von Hörern beurteilt werden. Diesen Punkt greift Kehrein (2002) hypothetisch auf, wenn er formuliert, dass

[...] diskrete Emotionen reine Wahrnehmungskategorien sind, die anhand der Beobachtung des komplexen Ausdrucksverhaltens (unter Berücksichtigung aller zur Verfügung stehender Signalisierungssysteme) in einer bestimmten Situation/Interaktion identifiziert werden. (Kehrein 2002:141)

Hierbei interessiert vor allem die Frage danach, inwieweit das auditive Signalisierungssystem die Beurteilung von Sprechaufnahmen bezüglich der Identifikation von diskreten Emotionen beeinflussen kann.

Ladd et al. (1985) führen dazu drei miteinander in Beziehung stehende Experimente durch, bei denen Hörer die Erregung von aufgezeichneten kurzen Äußerungen beurteilen sollen. Dabei werden systematisch und unabhängig voneinander die Stimmqualität, der Intonationsverlauf und die Grundfrequenz durch digitale Resyntheseverfahren natürlich gesprochener Äußerungen manipuliert. Im Ergebnis halten sie fest:

Analyses of variance of the results showed that range and contour, and less clearly range and voice quality, had independent effects on the way the utterances were judged. [...] Finally, analysis of the results suggests that differences of  $F_0$  range, as is commonly assumed, have continuous rather than categorical effects on affective judgments. (Ladd et al. 1985:435)

In einer Arbeit von Banse & Scherer (1996) belegen die Autoren, dass Hörer mit mehr oder weniger großer Genauigkeit emotionale Attribute eines Sprechers anhand der Sprechweise inferieren können. Um dies nachzuweisen, ließen sie Schauspieler 14 verschiedene Emotionen darstellen, die sich sowohl in ihrer Intensität als auch in ihrer Valenz unterschieden, und ließen die Fassungen im Anschluss von Hörern beurteilen. Die Daten suggerieren, dass vokale Parameter nicht nur einen Index für den Grad der Intensität einzelner Emotionen darstellen, sondern darüber hinaus auch einen Hinweis in Bezug auf die Differenzierung von Valenz- und Qualitätsaspekten von Emotionen geben (Banse & Scherer 1996:614). Eine Vielzahl von Arbeiten betrachtet in ihren Experimenten allerdings nur eine geringere Zahl von Emotionen. Tabelle 5 soll die Ergebnisse für das Diskriminieren einzelner Emotionen in Prozent zweier empirischer Studien im Vergleich abbilden.

Study	Fear	Disgust	Joy	Sadness	Anger
<b>van Bezooijen (1984)</b>	58	49	72	67	74
<b>Scherer et al. (1991)</b>	52	28	59	72	68

**Tab. 5: Vergleich der Prozentzahlen für das Diskriminieren einzelner Emotionen in zwei empirischen Studien. (Banse & Scherer 1996:615).**

Auch Bergmann et al. (1988) haben in vier Experimenten Attributionsprozesse untersucht, indem sie verschiedene manipulierte Äußerungen von separaten Beurteilergruppen bezüglich der emotionalen Eindruckswirkung eines Sprechers einschätzen ließen. Sie beschreiben:

Hochsignifikante Haupteffekte (bei nahezu völligem Fehlen von Interaktionseffekten) zeigen, daß die meisten der untersuchten Parameter einen deutlichen und spezifischen Einfluß auf die Hörerattribution von Sprecherzuständen haben. Besonders starke Attributionseffekte zeigen die Bandbreite der Stimm-Grundfrequenz-Variation (breit = hohe Aktivierung, schmal = Traurigkeit), Intensität (hoch = negative Effekte, aggressiver Hörerzustand), und Segmentdauer (kurz = eher freudig, lang = eher traurig). (Bergmann et al. 1988:167)

Durch die Möglichkeit digitaler Resynthese gesprochener Sprache ergeben sich auch für die empirischen Untersuchungen von Attributionsprozessen neue Möglichkeiten. So können beispielsweise durch unabhängige und faktorielle Variationen mehrerer akustischer Parameter Sprachäußerungen systematisch variiert werden. Umfangreiche theoretische und empirische Ausführungen zu Veränderlichkeiten der Grundfrequenz und deren Auswirkungen auf die Perception legen 't Hart et al. (1990) mit ihrem IPO Approach (experimental-phonetic approach to intonation) vor. Mit dem Fokus auf Perzeptionsprozessen versuchen die Autoren ein Modell der Datenreduktion zu entwickeln, welches Perception als Filter nutzt, um Daten, die nicht perzeptionsrelevant sind, auch als nicht kommunikationsrelevant auszuschließen. Bei dieser Methode wird der  $F_0$ -Verlauf soweit simplifiziert, dass er nur noch die perzeptionsrelevanten Bewegungen der Äußerung abbildet. Solche Konturen können, obwohl sie sich auditiv von den Originalfassungen unterscheiden, dennoch als melodisches Equivalent verstanden werden.

Eine ausschließlich empirische Untersuchung stellt die Studie von Burkhardt & Sendlmeier (1999) dar, worin die Autoren den Einfluss akustischer Parameter auf den emotionalen Ausdruck von Sprachäußerungen am Beispiel von Freude untersuchen. Dazu wird ein neutraler Satz hinsichtlich der Parameter Lage der

Formanten, Form der Grundfrequenzkontur und der Dauer der Laute systematisch durch PSOLA- und LPC-Verfahren variiert. Die Lage der Formanten wurde auf zwei Arten manipuliert: a) der erste Formant wurde um 6% und der zweite um 4% angehoben und b) die ersten fünf Formanten wurden um 5% erhöht. Beide Manipulationen der Formantlagen werden als freudiger beurteilt, wobei sich kein signifikanter Unterschied zwischen den beiden Manipulationen ergibt. Die erste Version wurde allerdings tendenziell etwas besser bewertet. Vermutet wird, dass sich die freudige Sprechweise durch mehr Energie in den höheren Frequenzbereichen auszeichnet (Klasmeyer & Sendlmeier 1999). In Bezug auf die  $F_0$ -Kontur zogen um 25% angehobene Satzbetonungen die höchste Bewertung für eine als freudig empfundene Sprechweise nach sich. Bei der Manipulation der Lautdauern wurden drei phrasenbetonte Silben um 24% verlängert und alle anderen Silben um 20% verkürzt, was ebenfalls zu einer als freudiger empfundenen Bewertung führte. Insgesamt ergab ein Hörversuch signifikante Effekte aller drei Parameter sowie eine Interaktion zwischen Grundfrequenzkonturen und Lautdauern.

Mit der Fragestellung, ob das Identifizieren und Diskriminieren von affektiver und linguistischer Prosodie alters- und geschlechtsabhängig ist, haben sich Raithel & Hielscher-Fastabend (2004) beschäftigt. Sie kommen dabei zu dem Ergebnis, dass affektive Prosodie von jüngeren Versuchsteilnehmern signifikant besser erkannt wird, wohingegen sich aber keine geschlechtsspezifischen Effekte nachweisen lassen. Diese Ergebnisse stehen im Kontrast zu den Daten von Breitenstein et al. (1996).

Schröder (2000 & 2004) beschäftigt sich im Rahmen verschiedener Arbeiten mit der Perzeption von Interjektionen. Er lässt von Schauspielern Interjektionen (Affect bursts) für zehn verschiedene Emotionen produzieren und unterzieht diese einem Perzeptionstest. Die Erkennungsraten liegen mit durchschnittlichen 81,1% für alle Emotionen weit über den Erkennungsraten, die Banse & Scherer (s.o.) für größere Äußerungseinheiten ermittelt haben. Es ist daher anzunehmen, dass sich einige Emotionen (z.B. Ekel) durch kurze und prägnante prosodische Sequenzen deutlicher ausdrücken und wahrnehmen lassen als durch größere Sprechereinheiten (Schröder 2004:56).

Auch Unterschiede in der Stimmqualität lassen verschiedene Rückschlüsse auf eine Sprecherbefindlichkeit zu. So zeigen Gobl & Chasaide (2000) in einer Untersuchung, dass allein durch Veränderungen der Stimmqualität unterschiedli-

che affektive Färbungen eines Sprechers evoziert werden können. In einem sehr kontrollierten Syntheseverfahren werden sieben Stimmqualitäten kreiert und einem Perzeptionstest unterzogen, wobei die Stimmqualitäten auf einer 7-er Skala in einem semantischen Differenzial als „gestresst/entspannt“ bzw. „zufrieden/ärgerlich“ eingeschätzt werden sollen. Im Ergebnis lassen sich die verschiedenen Stimmqualitäten in zwei Gruppen zusammenfassen: a) eine gespannte, raue Stimme, die als eher aggressiv empfunden wird und b) eine behauchte, knarrende Stimme, die als eher entspannt und unaggressiv eingeschätzt wird. Die Autoren weisen aber darauf hin, dass alle Ergebnisse vorsichtig interpretiert werden müssen, und dass sich aus der Stimmqualität allein keine eindeutigen Rückschlüsse auf eine konkrete Emotion ziehen lassen.

Cauldwell (2000) kann in seiner Arbeit nachweisen, dass Äußerungen, je nach dem, ob sie in einem Kontext oder isoliert präsentiert werden, unterschiedlich interpretiert werden. Zwei kurze Äußerungen (W-Fragen mit terminalem Melodiemuster) wurden isoliert wahrgenommen als ärgerlich empfunden, wohingegen dieselben Äußerungen in einen Kontext eingebettet als völlig neutral eingeschätzt wurden. Ähnliche Phänomene beschreiben Greasley et al. (2000), die feststellten, dass positive und neutrale Wörter in einem negativen Kontext negativer bewertet werden als in vergleichbarer Situation, aber isoliert präsentiert. Sie führen dies auf unterschiedliche linguistische und paralinguistische Funktionen zurück.

### 3.4 ZUSAMMENFASSUNG

Die aufgeführten Arbeiten geben einen Überblick über den Zusammenhang stimmlicher Parameter und deren Identifikation bzw. Perzeption im Kontext emotionalisierter Sprechweisen. Dabei liegen der Planung und Durchführung der Experimente methodisch sehr unterschiedliche Ansätze zu Grunde. Die einfachste und hinsichtlich der Datengewinnung am wenigsten aufwendige Methode ist die Simulation von verschiedenen Emotionen durch Schauspieler (Abadjeva et al. 1993, Banse & Scherer 1996, Kienast 2002, Paeschke 2003, Kranich 2003). Dies bietet zum einen den Vorteil, dass durch konstante Laborbedingungen die Sprachaufnahmen qualitativ hochwertig sind und zum anderen, dass die Aufnahmen gut vergleichbar sind, weil die Analyse identischer

Texte eine Trennung von linguistischen und paralinguistischen Funktionen der Prosodie ermöglicht. Allerdings liegt der Nachteil von simulierten Emotionen darin, dass Unterschiede zwischen echten und unechten Emotionen durchaus erkannt werden. So konnte Ekman (1984) herausfinden, dass Versuchspersonen den Gesichtsausdruck überzeugend simulierter Emotionen dennoch erkennen, was sich mit Sicherheit auch auf den Stimmausdruck übertragen lässt.

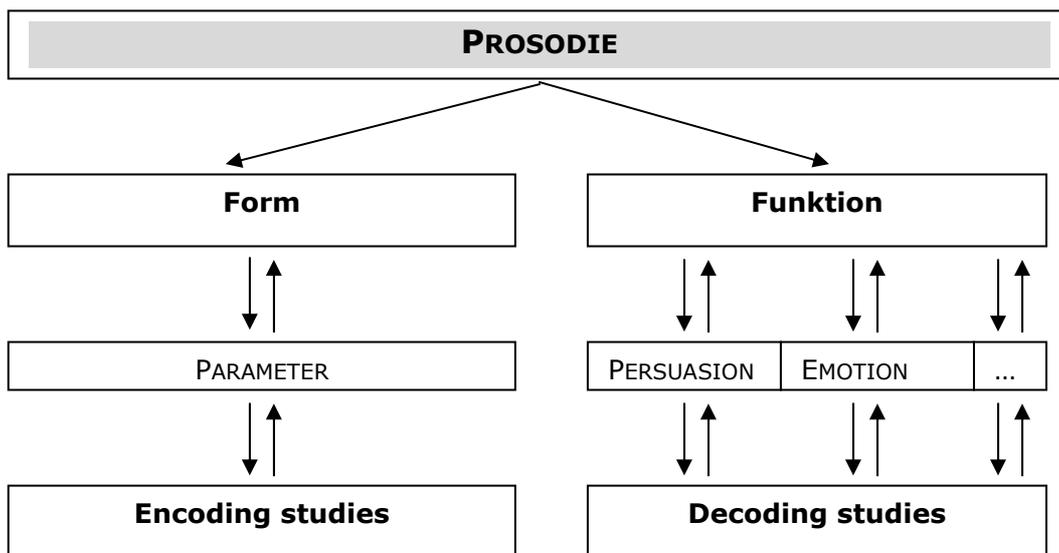
Für die Verwendung authentischer Emotionen kann die Arbeit von Williams & Stevens (1972) angeführt werden. Die Autoren waren in der seltenen Lage, sowohl echte, spontane Emotionen als auch nachträglich von Schauspielern reproduzierte Emotionen zu untersuchen. Da es die aufwendigste Methode zu sein scheint (auf diese Weise ein umfangreiches Korpus zu erstellen, was verschiedene Emotionen abdeckt und dennoch eine Vergleichbarkeit der Daten bei guter Qualität gewährleistet, gestaltet sich schwierig), hat sich diese Methode nicht wirklich in den wissenschaftlichen Untersuchungen bewährt.

Eine andere Möglichkeit, Emotionen zu simulieren, wird in den Arbeiten von Fernandez & Picard (2000), Kehrein (2002) oder Amir et al. (2000) gewählt. Hier werden die Probanden durch komplexe Versuchsanleitungen in Stress versetzt und müssen dabei verschiedene Aufgaben lösen. Die dabei entstandenen Aufnahmen werden aufgezeichnet und spiegeln emotionalisierte Sprechweisen wieder, die anschließend unter verschiedenen Gesichtspunkten analysiert werden können. Der Nachteil dieser Art der Datengewinnung besteht darin, dass sich kein besonders breites Spektrum an Emotionen erzeugen lässt. Meistens handelt es sich um verärgerte Sprechweisen im Vergleich zu einer neutralen Variante.

Neuere Arbeiten versuchen über Sprachsyntheseverfahren Emotionen künstlich zu erzeugen bzw. so zu manipulieren (Burkhardt & Sendlmeier 1999, Schröder 2004), dass sich auf diese Weise konkrete Aussagen über einzelne Parameter und deren Zusammenwirken ableiten lassen. Die Methodenvielfalt weist schon auf die Verschiedenartigkeit der einzelnen Untersuchungen hin.

Eine Möglichkeit, die Forschungsliteratur zu systematisieren, bietet die von Scherer (2000a) vorgeschlagene Differenzierung von Encoding- bzw. Decoding studies, in Anlehnung an das Brunswicksche Linsenmodell (s.o.), die sich entweder aus der Sprecherperspektive mit dem Verschlüsseln von Emotionen durch stimmliche bzw. sprecherische Parameter beschäftigen, oder aus der Hörerperspektive mit der Perzeption einzelner Emotionen anhand konkreter

Sprachaufzeichnungen. Diese Einteilung wiederum lässt sich gut auf die von Neuber (2002:51f) postulierte Dichotomie der Prosodie in eine Form- bzw. Funktionsseite (vgl. Kap. 2.3) übertragen und bildet somit abschließend einen exzellenten Brückenschlag zwischen den theoretischen Ausführungen der einzelnen Kapitel. In Abbildung 8 ist diese Synthese schematisch zusammengefasst.



**Abb. 8: Zusammenhang zwischen der Dichotomie des Prosodiebegriffs und einer Systematik zur Einordnung der Literatur.**

## 4 REZEPTIONSDETERMINANTEN UND PROSODIE

### 4.1 EXTERNE REZEPTIONSDETERMINANTE: SPRECHAUSDRUCK

Prosodische Mittel geben nicht nur in lingualer, sondern auch in paralingualer Verwendung Auskunft über Merkmale einer sprechenden Person. Unter phonetischen Gesichtspunkten lassen sich dabei zwei Kategorien voneinander unterscheiden, eine (a) physiognomische (individuumsspezifische) Kategorie, die sich angeborenen oder gewohnheitsmäßigen Gestaltungsweisen im Sprechen zuwendet und (b) eine pathognomische (situationsspezifische) Kategorie, die den überindividuellen Ausdrucksgehalt einer Äußerung fokussiert (Anders 2001a:27f, Bose 2001:266, Neuber 2002:35).

Der Deutung und Bewertung des pathognomischen Ausdrucks liegen dabei [...] die Beurteilungen individuumsspezifischer Merkmale der Stimme und Sprechweise des Sprechers und andererseits die Bewertungen vergleichbarer Ausdruckserscheinungen der eigenen Sprechweise und der anderer Sprecher in ähnlichen Situationen zugrunde. (Anders 2001a:28)

Im Zusammenhang mit der Beschreibung solcher soziophonetischer Funktionen und deren thematischer Abgrenzung werden oftmals biophonetische oder humanphonetische Arbeiten herangezogen, die aus dem Blickwinkel der Verhaltensforschung heraus davon ausgehen, dass die Äußerung eines Sprechers neben dem geplanten Sprechausdruck auch einen elementaren Stimmausdruck besitzt, der sich insbesondere auf die angeborenen stimmlichen Ausdrucksformen zur Charakterisierung von Basisemotionen wie Wut, Ekel, Angst und Freude bezieht. Dieser Sachverhalt wurde verschiedentlich in der Literatur thematisiert (Trojan 1952, 1975, Bastian 1986). Tembrock (1977, 1996) hat zu diesem Aspekt aus der Tierverhaltensforschung heraus wichtige Anstöße gegeben, indem er zahlreiche Ähnlichkeiten zwischen stimmlichen Ausdrucksmerkmalen vieler Tierarten (besonders der Säugetiere) und dem Menschen aufzeigt.

Der Stimmausdruck kann aber auch bewusst gesteuert und gestaltet werden. Dies geschieht z.B. auf der Bühne, wenn der Stimmausdruck durch die schauspielerische Gestaltung im Sinne einer bestimmten Äußerungsabsicht modifiziert und variiert wird. Und auch im Fernsehen werden in Werbespots konkrete stimmliche und sprecherische Ausdrucksmuster bewusst zur Gestaltung eingesetzt, um eine avisierte Äußerungsabsicht zu evozieren. Dabei gelingt es aber

nicht, allgemeingültige Kriterien für einzelne Ausdrucksqualitäten zu benennen, weil sowohl pathognomische als auch physiognomische Sprechausdrucks- und Stimmausdrucksmerkmale ständigen Wandlungen unterworfen sind, die darauf zurückzuführen sind, dass sich die persönliche und die soziale Identität eines Sprechers (und auch eines Hörers) ständig weiterentwickeln (Anders 2001a:28). So beschreiben Eckert & Laver (1994:35f) beispielsweise die „Kleine-Mädchen-Stimme“, anhand eines Filmes aus den 40er-Jahren, indem eine Mittvierzigerin in der Rolle der Gastgeberin als Frau eines Generaldirektors mit bewusst mädchenhafter Stimme spricht, um auf diese Weise eine untergeordnete Rolle zu signalisieren, auf welche Männer wiederum mit übergeordnetem Verhalten reagieren. Solche Klischees würden in aktuellen Filmszenen ihre Wirkung verfehlen und eher Lächerlichkeit hervorrufen.

Es ist heute gesellschaftlich eher akzeptabel, nicht so weiblich zu klingen. Eine tiefere Stimme wird mit Reife und Autorität in Verbindung gebracht. (Flensburger Tageblatt 14. 4. 1993, zitiert nach: Eckert & Laver 1994:36)

Eine Studie von Seidner & Büttner (1998) untersucht die Ästhetik heiserer Singstimmen durch verschiedene Hörergruppen (Phoniatler/Logopäden, Patienten, Studenten der Gesangspädagogik, Studenten nichtmusikalischer Fachrichtungen und Fußballspieler). Im Ergebnis finden die Autoren heraus, dass als heiser eingeschätzte Stimmen auch überwiegend als krank empfunden werden. In allen Hörergruppen gefielen aber unter ästhetischen Gesichtspunkten am meisten die Stimmen, die zwar als gesund, aber zugleich auch als leicht heiser eingeschätzt wurden. Die heisere Stimme stellt in der Popmusik ein künstlerisches Ausdrucksphänomen dar, das

[...] zu einem Soundkonzept im Sinne eines Markenzeichens gehört und die Stimme als Teil dieses Sounds vor allem starke Emotionen vermittelt. Sie soll Vitalität, Energie und auch Protest gegen Herkömmliches ausdrücken und damit beeindrucken. (Seidner & Büttner 1998:75)

Wandlungen hinsichtlich der Bewertung von stimmlichen und sprecherischen Ausdrucksmerkmalen sind u.a. zurückzuführen auf einen veränderten Zeitstil oder einen veränderten Kunstgeschmack (Götttert 1998). Daher müssen Wirkungsanalysen immer vor einem konkreten Hintergrund durchgeführt werden, der dem Zeitstil Rechnung trägt. Anders (2001a:11) legt, diesen Umständen eingedenk, ihren Untersuchungen einen Wirkungsbegriff zugrunde, der Wirkung als generelle Akzeptanz einer sprecherischen Fassung definiert und damit in enger Verbindung zum wissenschaftlichen Verständnis des Terminus in der

Sprechwirkungsforschung steht, wobei folgende begriffliche Präzision vorgenommen wird:

Ein für diesen Zweck abgeleiteter stimmlicher Ausdruck wird dann aber nicht nur nach kommunikativer Adäquatheit und Korrektheit, sondern auch nach dem Grad an Natürlichkeit oder Stilisierung, nach der Echtheit, der Anschaulichkeit, der schöpferischen Gestaltung und dem Grad der Wirkung auf den Hörer beurteilt. (Stock & Suttner 1991:69)

Um genauer untersuchen zu können, worauf die einzelnen Wirkungen des Stimm- und Sprechausdrucks zurückzuführen sind, schlägt Anders (2001a:28f) eine Differenzierung zwischen Sprechausdruck und Prosodie vor. Demnach stützt sich der Eindruck des Hörers immer auf erfassbare und konkret beschreibbare Äußerungsmerkmale der segmentalen und suprasegmentalen Struktur des Gesprochenen, d.h. um einer Äußerung den jeweiligen Sprechausdruck zumessen zu können, nimmt der Hörer eine ganz bestimmte Sprecherbefindlichkeit wahr. Somit liefern prosodische Merkmale dem Hörer auditiv wahrnehmbare Hinweise, anhand derer er auf pathognomische bzw. physiognomische Sprechereigenheiten schließen kann (vgl. Kap. 3.2). Prosodische Merkmale besitzen aber in der Koppelung an gesprochene Sprache keinen Ausdruckswert per se, sondern sie werden vom Hörer in der Phase der Rezeption als Anhaltspunkt für den Sprechausdruck genommen. Auf diese Weise unterliegt die Beurteilung des Sprechausdrucks als Interpretationsakt möglichen Irrtümern und subjektiven Einzelurteilen der Hörer. Daher erfordert der Umgang mit empirisch gewonnenen Daten eine methodische Absicherung dahingehend, dass den Auswertungen ein unverwechselbarer Beurteilungsmaßstab zugrunde gelegt wird, der Rückschlüsse auf gültige Hörerurteile zulässt (s.u.).

Um die Anzahl der prosodischen Mittel, die dem Sprechausdruck zugeordnet werden, finden sich in der Literatur differierende Auffassungen. Eine frühe Arbeit von Winckel (1960) nimmt zur Beschreibung des Stimm- und Sprachklanges anhand des Spektrums eine Unterscheidung in eine ästhetische Qualitätsstimme (z.B. guter Sprecher oder Sänger) und eine semantische Qualitätsstimme (diejenige, die am Telefon am besten verstanden wird) vor. Je nachdem, an welche Funktion eine Stimme gebunden ist, kommt dabei speziellen Spektren eine gesonderte Bedeutung zu. Nach Winckel würde man einen Fortschritt erreichen, wenn man

[...] eine Aufteilung in semantische und ästhetische Information vornehmen würde, weil man dann in dem ersten Faktor quantitative Angaben für Anomalien von Sprache und Stimme erhalten und in dem zweiten Faktor Personalkonstanten ermitteln könnte. In dieser Sicht erweist es sich dann, daß das Klangspektrum mehr ästhetischen als semantischen Wert hat [...]. (Winckel 1960:129)

Im Zusammenhang mit der Beurteilung wirkungsvoller Stimmen in der Werbung stehen vor allem ästhetische Stimmqualitäten im Vordergrund. Welche Merkmale dabei eine Stimme zu einer ästhetischen Stimme machen, wird Gegenstand der empirischen Untersuchung dieser Arbeit sein (vgl. Kap. 5 und Kap. 6).

Verschiedene Arbeiten, die sich sowohl theoretisch als auch empirisch mit prosodischen Merkmalen des Sprechausdrucks unter verschiedenen Fragestellungen beschäftigt haben (Anders 2001a, Bose 2001, 2003, Kranich 2003), schlagen einen Kanon an Merkmalen vor, die für das Untersuchen und Erfassen des Sprechausdrucks von Bedeutung sind. Diese Übersichten zeichnen sich allerdings oftmals durch Merkmale aus, die in der Literatur nicht eindeutig definiert sind und die sich daher schwierig bestimmen lassen (Merkmale wie z.B. faukale Distanz oder Sprechspannung). In dieser Arbeit wird deshalb auf einen Kanon von Fährmann (1967) zurückgegriffen, der, weil er in den sechziger Jahren entstand, zwar etwas weniger detailliert ist, aber alle Parameter der Stimme und Sprechweise nennt, die für die eigene Untersuchung von Bedeutung sein werden (Tab. 6). Sein Merkmalskatalog ist charakterisiert durch zwei Kategorien, denen er einzelne Parameter zuschreibt, nämlich (a) die habituellen Stimmqualitäten, die sich auf die physiologischen Voraussetzungen des Sprechers beziehen und (b) die individuellen Verlaufsqualitäten, die sich auf die sprecherische Realisierung einer Äußerung beziehen. Die von Fährmann ebenfalls angeführten (c) formalen Kriterien wie Sprechstil und Sinnform werden hier außer Acht gelassen, weil sie keine Bedeutung für die Fragestellung der eigenen Untersuchung haben, sie gewinnen eher in sprechkünstlerischen Kontexten oder unter sprechästhetischen Fragestellungen an Bedeutung. In Tabelle 6 sind diejenigen Kriterien aus Fährmanns Katalog, die für den weiteren Verlauf der Arbeit relevant sein werden, mit ihren jeweils zugehörigen Parametern überblicksartig aufgeführt und in ihren wesentlichen Punkten erläutert.

Der Merkmalskatalog (Tab. 6) bietet eine Systematik zur auditiven Sprechausdrucksbeschreibung. Demgegenüber stehen instrumentelle Korrelate (teilweise komplexe Messungen) im akustischen Signal. Terminologisch orientieren sich

dabei vor allem sprechwissenschaftliche Arbeiten (Geißner 1981, 1985; Krech et al. 1991, Bose 2001, Kranich 2003) an den angeführten Beschreibungssystemen.

---

Habituelle Stimmqualitäten	<ul style="list-style-type: none"> <li>▪ Tonhöhe (Sprechstimmlage)</li> <li>▪ Lautstärke</li> <li>▪ Stimmfülle (Volumen)</li> <li>▪ Klangfarbe (Timbre)</li> </ul>
Individuelle Verlaufsqualitäten	<ul style="list-style-type: none"> <li>▪ Sprechtempo</li> <li>▪ Akzentuierung (melodisch, dynamisch,agogisch [temporal])</li> <li>▪ Artikulation</li> </ul>

---

**Tab. 6: Parameter der Stimme und Sprechweise nach Fährmann (1967:28ff).**

Schwierigkeiten im Erstellen eines allgemeingültigen Merkmalskataloges zur Beurteilung von Stimme und Sprechweise lassen sich vor allem auf unterschiedliche Beschreibungssysteme innerhalb benachbarter Disziplinen zurückführen. So werden Termini aus sprechwissenschaftlichen Arbeiten (Neuber 2002, Kranich 2003, Bose 2001, 2003) in Arbeiten, die sich an Beschreibungssystemen zur Diagnostik gestörter Stimmen und Sprache orientieren (Wirth 1995, Pfau/Streubel 1982, Wendler/Seidner 1977, Luchsinger/Arnold 1970) und Arbeiten, denen akustische Fragestellungen zu Grunde liegen (Klasmeyer/Sendlmeier 2000, Neppert 1999) in ihrer begrifflichen Präzision oftmals sehr unterschiedlich verwendet. Die im Folgenden genauer ausgeführten Parameter stellen eine Auswahl in Bezug auf ihre Bedeutung für die sich anschließende Untersuchung dar und sind verschiedenen Beschreibungssystemen entnommen. Unter dem Gesichtspunkt der Charakterisierung und Abgrenzung verschiedener Sprechfassungen voneinander bilden sie gleichzeitig die theoretische Grundlage für den experimentellen Teil der Arbeit, der in Kapitel 5 und 6 beschrieben wird.

## **Tonhöhe (Sprechstimmlage, Grundfrequenz)**

Terhardt beschreibt Tonhöhe in seinen Ausführungen als

[...] diejenige Hörempfindung [...], welche den Aspekt der Periodizität und der Periodendauer der Schallsignale repräsentiert. (Terhardt 1998:307)

Damit wird eine Verbindung zwischen apparativen und auditiven Beschreibungsmethoden skizziert, die aber auf keinem streng linearen Zusammenhang beider Beschreibungssysteme beruht. So argumentiert Neppert (1999:31) beispielsweise damit, dass Hörempfindungen u.a. auch von Spektralmerkmalen der Stimme (z.B. Helligkeit) abhängig seien. Als akustisches Korrelat der Tonhöhe wird die Grundfrequenz ( $F_0$ ) gemessen. Für die Perzeption akustischer Phänomene, ein Bereich mit dem sich die Psychoakustik (vgl. Zwicker 1982) beschäftigt, ist charakteristisch, dass es keinen linearen Zusammenhang zwischen einem auditiv wahrgenommenen Tonhöhenanstieg und steigender Grundfrequenz gibt. Durch die logarithmische Skalierung des Frequenzmaßes Hertz (Hz) und durch das ebenfalls einem logarithmischen Kurvenverlauf ähnelnden Empfindungsvermögen des Menschen für Tonhöhen ergibt sich eine besonders sensible Tonhöhendifferenzierung in bestimmten Frequenzbereichen. So z.B. für den Bereich zwischen 100 Hz – 250 Hz, dem Intervall, innerhalb dessen gewöhnlich die Stimmlippenschwingungen für die einzelnen Stimmgattungen angesiedelt sind (Kranich 2003:20). Um die fehlende Linearität der logarithmischen Kurven apparativ und auditiv gewonnener Daten auszugleichen, wird die Differenz der Tonhöhenbewegung zwischen zwei Messpunkten in Halbtonschritten angegeben. Der Vorteil dieser Darstellung liegt in der linearen Umrechnung von Tonhöhen, woraus sich eine bessere Lesbarkeit der Daten ergibt, was vor allem den unmittelbaren Vergleich von Tonhöhendifferenzen zwischen höheren und niedrigeren Frequenzunterschieden ermöglicht.

Als mittlere Sprechstimmlage wird die durchschnittlich wahrgenommene Sprechtonhöhe bezeichnet, um die sich Tonhöhenvariationen ca. innerhalb einer Quarte bewegen (Bose 2001:279, Wendler et al. 1996:99, Pahn 1988:33). Dieser Bereich wird durch die Indifferenzlage begrenzt. Sie ist:

die physiologische Norm der Sprechstimmlage; sie wird vor allem vom Körperbau, speziell von den Dimensionen der Stimmlippen bestimmt und hängt gesetzmäßig mit der Stimmgattung zusammen. (Stock 1998:196)

Für die einzelnen Stimmgattungen ergeben sich für die Sprechstimme durchschnittliche Sprechstimmlagen, die Wirth (Tab. 7) in folgenden Bereichen einordnet:

<b>Baß</b>	G – A	98.99 – 110.00 HZ
<b>Tenor</b>	c	130.81 Hz
<b>Alt</b>	unterhalb gis	207.65 Hz
<b>Sopran</b>	h – c <sup>1</sup>	246.94 – 261.62 Hz

**Tab. 7: Mittlere Sprechstimmlage für die unterschiedlichen Stimmgattungen nach Wirth (1995:123).**

### **Lautstärke (Lautheit, Intensität)**

Terminologisch ergeben sich auch in Bezug auf die Intensität Differenzen zwischen dem Begriff Lautstärke, der im physikalischen Sinn (gemessen in dB) eine „flächenbezogene Schalleistung“ (Terhardt 1998:271) bezeichnet. Und dem Begriff der Lautheit, der als physiologische Größe (gemessen in Phon) nach Neppert (1999:29) die an einem Schall wahrgenommene Lautstärkeempfindung als subjektives Korrelat beschreibt. Generell lässt sich festhalten, dass die Lautheit mit physikalisch messbarer Schallamplitude wächst, dass sich aber ähnlich wie bei der Perzeption der Tonhöhe auch hier keine lineare Beziehung herstellen lässt. Vieregge (1996:161) greift diesen Punkt auf und fügt ergänzend hinzu, dass die Wahrnehmung von Lautstärkephänomenen neben der messbaren Intensität auch von anderen Faktoren wie dem subjektiven Gesamteindruck der Sprechstimme, von der Dauer des Sprechereignisses, von den Vokalqualitäten und von der auditiv-visuell wahrnehmbaren Sprechanstrengung abhängen.

### **Stimmfülle (Klangfülle, Volumen)**

Die Klangfülle einer Stimme ist nach Seidner/Wendler (1982:152) physiologisch durch Regelmäßigkeiten im Schwingungsablauf der Stimmlippen und durch Resonanzeigenschaften des Ansatzrohres, die die Teiltonstruktur des Spektrums bestimmen, gekennzeichnet. Demnach lassen sich klangvolle (resonanzreiche, teiltonreiche, weittragende, volle, kräftige, volumenreiche)

Stimmen von klangarmen (resonanzarmen, teiltonarmen, zarten, dünnen, volumenarmen, klanggesättigten) Stimmen unterscheiden. Dabei muss kritisch angemerkt werden, dass durch die fehlende begriffliche Präzision, z.B. bei der Verwendung der Termini teiltonreich gegenüber teiltonarm, falsche Assoziationen hervorgerufen werden. Gemeint ist nicht, wie die Termini aber vermuten lassen, die Quantität der einzelnen Teiltöne, sondern die Qualität und Stärke der Teiltöne innerhalb des Spektrums. Das akustische Korrelat der Stimmfülle betrachtet das Teiltonspektrum bis in hohe Frequenzen. Sind die höheren Formanten, z.B.  $F_4$  und  $F_5$ , immer noch hoch verstärkt, ist dies Indiz einer gut schließenden Glottis und einer damit einhergehenden resonanzreichen Stimme. Eine Ursache für die Stimmfülle liegt also an der Quelle, der Glottis selbst. Je nach dem Schließungsgrad der Glottis und den Spannungsverhältnissen an der Glottis verändert sich das Teiltonspektrum charakteristisch, was sich wiederum in unterschiedlichen Stimmqualitäten wie Modalstimme, verhauchte oder verknarrte Stimme (s.u.) äußert.

Eine andere Ursache der Stimmfülle liegt in der Resonanz des Ansatzrohres, welches sich oberhalb der Quelle (Glottis) befindet und die drei Hohlräume Rachen-, Mund- und Nasenraum umfasst. Durch Modifikationen des Luftstroms in einzelnen Resonanzräumen des Ansatzrohres werden bestimmte Formanten verstärkt, die Einfluss auf die Stimmfülle haben (z.B. Nasalresonanz). Ausschlaggebend für resonanzreiche Stimmen sind dabei zum einen eine hohe Übereinstimmung der Resonanzfrequenzen mit den Teiltönen und zum anderen der Härtegrad der Wandungen des Ansatzrohres. Je härter die Wandungen des Ansatzrohres sind, desto weniger Energie kann verloren gehen (Sundberg 1997:32ff, Wirth 1995:88f).

Das Problem der Objektivierung von Klangunterschieden durch Messverfahren ist bisher nicht befriedigend gelöst. Auch für die auditiven Beurteilungen, die in der Praxis immer noch überwiegen, gibt es in der Fachliteratur keine einheitliche Bewertungsskala (vgl. Wirth 1995, Friedrich/Bigenzahn 1995, Seidner/Wendler 1982, Pfau/ Streubel 1982). Dennoch kann davon ausgegangen werden, dass professionelle Hörer zu ähnlichen Urteilen in der Einschätzung von Klangmerkmalen der Stimme kommen,<sup>8</sup> was auch die Ergebnisse der eigenen Untersuchung (s.u.) zeigen.

---

<sup>8</sup> Die Ähnlichkeit von Hörurteilen lässt sich auch auf das Training und die Hörgewohnheiten innerhalb einzelner Schulen zurückführen, wo das Hören zumeist impressionistisch geschult wird.

Die Termini Klangfülle und Klangfarbe sind eng miteinander verbunden, bedingen einander und lassen sich nicht präzise voneinander abtrennen. Daher stellt der nächste Punkt weniger eine neue begriffliche Kategorie als vielmehr einen separaten Schwerpunkt dar.

### **Klangfarbe (Timbre, Stimmklang, Stimmqualität)**

Auch die Beschreibung der Klangfarbe lässt sich auf Modifikationen von Resonanzverhältnissen im Ansatzrohr oder auf Veränderungen an der Glottis zurückführen.

Die Darstellung der Klangfarbe oder des Timbres von Stimmen durch objektive Bestimmungen (z.B. mittels Spektrografie) ist Gegenstand verschiedener Untersuchungen (Baken/Orlikoff 2000, Neppert 1999, Anders 1997, 2000), dennoch ist die Datenlage unzureichend, um gesicherte Aussagen ableiten zu können, die es erlauben, akustische Sprechstimmerkmale zu katalogisieren (Stock 1996:74, Kranich 2003:33f). Gemeinsamkeiten bestehen allerdings in der Annahme, dass auch die Klangfarbe mit dem Teiltonspektrum zusammenhängt, welches nach Seidner/Wendler (1982:30) durch verschiedene Ansatzrohrformungen erzeugt wird. Neuber (2002) geht dabei von einem Merkmalskomplex aus.

Der Klangeindruck als Sinnesempfindung wird von der Parameterkombination aus Frequenzen und Amplituden sowie dem Frequenzverhältnis der Teiltöne zum Grundton bestimmt. (Neuber 2002:80)

Der Stimmklang ist also akustisch durch seine Periodizität gekennzeichnet. Neuber (2002:70f) konnte in einer Untersuchung zeigen, dass Veränderungen des Stimmklanges eingesetzt werden, um emotionale Unterschiede in Sprechäußerungen zu markieren. So beschreibt er für den Vokal /a/ ausgesprochene Periodizität in einer freundlichen Äußerung, wohingegen in gleicher Position bei wütender Sprechweise Aperiodizität beobachtet werden konnte.

Die Wahrnehmung der Klangfarbe einer Stimme ist allerdings oft auch beeinflusst von der Tonhöhe, wobei höhere Stimmen meist heller wirken. Nach Wirth (1995:124) sind die Mittelwerte der Formantpositionen  $F_1 - F_4$  Ausdrucksindizes unterschiedlichen Timbres. Diese Annahme teilt auch Neppert (1999), indem er festhält:

Dunklere Klänge und Rauschschalle haben ihre stärkeren Teilschwingungen überwiegend im Bereich der niedrigen Frequenzen. [...] Bei helleren Schallen liegt der spektrale Energieschwerpunkt im Bereich der relativ höheren Frequenzen. [...] Bei mittlerer Helligkeit liegt der Energieschwerpunkt in mittleren Frequenzbereichen. (Neppert 1999:65f)

Neben dem Versuch den Stimmklang akustisch zu beschreiben, besteht auch das Bedürfnis, das Phänomen Klangfarbe auditiv präziser zu fassen. In Anlehnung an verschiedene Stimmeigenschaften von Instrumenten in der Musik (Reuter 2001) entstanden zahlreiche Inventare an Metaphern zur Darstellung des Timbres (Brunner 2000:151, Tillmann 1974:88ff, Luchsinger/Arnold 1970:217ff, Wendler/Seidner 1977:133f). Dabei sind die Bezeichnungen zur qualitativen Stimmklangbeschreibung unterschiedlichen Wirklichkeitsbereichen entnommen, wie z.B. weich, hart (Härtegrade), warm, kalt (Temperatur), hell, dunkel (farbliche Beschreibung). Obwohl die Vermutung nach großer subjektiver Streuung bezüglich der auditiven Bewertung von Timbreeigenschaften nahe liegt, argumentiert Kranich, dass von ähnlichen Wahrnehmungsmustern ausgegangen werden kann (2003:94), was die Ergebnisse der eigenen Untersuchung (Kap. 6.2) bestätigen.

Auch bezüglich der Beurteilung der Stimmqualität existieren verschiedene Inventare von Adjektiven, die von hauchig, reibend, rasselnd bis rau, knarrend oder gepresst reichen. Eine gute Klassifizierungsmöglichkeit stellt der RBH-Index von Anders (1997, 2000, 2001) dar. Anhand von Klangbeispielen wird das Gehör geschult, um den genauen Grad an Rauigkeit (irreguläre Stimmlipenschwingung), Behauchung (Glottisschlussinsuffizienz) und Heiserkeit (Kombination aus Rauigkeit und Behauchung) von Stimmen quantitativ möglichst einheitlich beurteilen zu können. Man darf dabei allerdings nicht außer Acht lassen, dass phoniatische Klassifikationen auf die pathologische Stimme zugeschnitten sind. Dennoch bietet dieser Kanon bei professionellen Hörern auch für die Beurteilung phoniatisch unauffälliger Stimmen eine geeignete Grundlage.

### **Sprechtempo (Sprechgeschwindigkeit, Pausen)**

Mit Sprechtempo wird die Realisierung bestimmter sprachlicher Einheiten (Laute/Silben/Wörter) in einer bestimmten Zeiteinheit ausgedrückt. In der Phonetik stellte sich die Zählung von Silben/Sekunde als vorteilhaft heraus (Meinhold 1995:45). Die auditive Wahrnehmung hängt allerdings auch von

zahlreichen anderen Faktoren wie Pausenlängen, Pausenhäufigkeiten, Pausenpositionen, von der Akzentuierungsstärke, der Artikulationspräzision und von Lautdehnungen ab (Bose 2001:279). Lehiste (1976) zeigt, dass auch Grundfrequenzverläufe die Perzeption der Dauer beeinflussen. In einem Versuch bekamen Hörer Stimuli je gleicher Länge, die aus zwei identisch langen Einheiten bestanden, vorgespielt. Bei gleich bleibender Grundfrequenz wurde jeweils der erste Teil als länger empfunden. Wenn aber auf dem zweiten Teil eine Grundfrequenzmodulation (steigend-fallend, fallend-steigend) stattfand, wurde der zweite Teil länger empfunden als der erste Teil der Äußerung.

Sowohl die Sprechgeschwindigkeit als auch die Pausensetzung schwanken im frei formulierenden Sprechen oft kleinräumig, da sie abhängig von der Sprechgliederung und der Sprechausdrucksgestaltung sind (Bose 2001:279). Für die eigene Untersuchung sind die Daten anderer Untersuchungen zur Pausensetzung und zum Sprechtempo nur marginal relevant, weil es sich um sehr kurze und vorformulierte Sprecheinheiten handelt.

### **Akzentuierung (Melodische Veränderlichkeit, Dehnungsakzente)**

Akzentuierungen stellen komplexe phonetische bzw. prosodische Phänomene dar, wobei durch Markierungen melodischer, dynamischer, temporaler oder gemischter Art, Elemente einer Äußerung (Silben, Wörter) vor ihrem Hintergrund kontrastiert werden. Hirschfeld (1995:56) und Benkwitz (2004:19) nennen Veränderungen der Tonhöhe, größere Lautstärke, geringeres Sprechtempo und größere Spannung und Deutlichkeit als typische Merkmale der Akzentuierung. Kranich (2003) baut diesen Gedanken aus:

Akzentuierungserscheinungen stellen typischerweise einen Komplex verschiedener suprasegmentaler Eigenschaften dar. Dabei steht meist ein bestimmtes prosodisches Element im Vordergrund [...]. Andere Suprasegmentalia verstärken indes die Kontrastwirkung. (Kranich 2003:29)

Da es sich bei den allermeisten Akzenten um Merkmalskomplexe handelt, weist Bose (2001:281) zu Recht auf die Problematik der Perzeption von Akzenten hin und spricht von relativen Bestimmungen, bei denen auditive Beurteilung und akustische Messung durchaus voneinander abweichen können.

Melodische Veränderlichkeiten und Dehnungsakzente stellen häufige Markierungsmöglichkeiten dar und haben sich bei Kranich (2003) und in der eigenen Untersuchung als aussagekräftige Parameter erwiesen. Physiologische Vor-

raussetzungen für Sprechtonhöhenbewegungen sind Veränderungen in der Stimmlippenspannung, verschiedene Einstellungen des Spannapparates der Kehlkopfmuskulatur und Veränderungen des subglottischen Drucks (Seidner/Wendler 1982:81). Diese Veränderungen lassen sich apparativ durch den  $F_0$ -Verlauf nachweisen oder durch eine auditive Einschätzung des Sprechmaterials. Als Klassifizierungsmethode für die Höreinschätzung wählt Kranich (2003:82ff) eine Differenzierung in quantitative und qualitative Akzente, um somit Abweichungen und Extrempositionen gezielt erfassen zu können. Dabei lässt er einschätzen: (a) wie häufig wechselt die Sprechtonhöhe (nicht/ selten – mittel – mehrfach) und (b) wie groß ist der Intervallsprung (klein – mittel – groß). Parallel dazu ergaben sich für die Dehnungsakzente die Fragen (a) wie viele Akzentstellen werden als Phonemdehnung wahrgenommen (kein/seltenes Auftreten – mittleres Auftreten – mehrfaches Auftreten) und (b) wie groß ist das Maß des Dehnungsgrades in Bezug zur vorherigen Silbe (keine/kaum wahrnehmbare Dehnung – mittlere Dehnung – starke Dehnung).

### **Artikulation (Artikulatorische Spezifika)**

Im vorliegenden Fall liegt der apparativen und der auditiven Analyse weniger die Norm der deutschen Standardaussprache zugrunde, als vielmehr der Aspekt der die Artikulation begleitenden zusätzlichen Merkmale wie Aspiration, Glottalisierung, Entstimmlichungen und Nasalierungserscheinungen, die quasi als prosodische Merkmale zum segmentalen Gehalt der Äußerung hinzukommen. Auch verkürzte oder gedehnte Lautdauern begleiten die Artikulation und sollen als Phänomene beschrieben werden.

Diese Ausführungen bilden gleichsam den Übergang von diesem Kapitel, in dem es primär um die externe Rezeptionsdeterminante des Sprechausdrucks ging, zum nächsten Kapitel, in dem es vordergründig um die interne Rezeptionsdeterminante und die damit verbundenen Beurteilungsmaßstäbe gültiger Hörerurteile gehen soll.

## 4.2 INTERNE REZEPTIONSDETERMINANTEN: HÖRMUSTER UND HÖRGEWOHNHEITEN

Verschiedene Arbeiten, die sich mit der Bewertung von Sprechausdrucksmerkmalen in Form von auditiven Analysen beschäftigen, sind auf das Problem der Objektivierbarkeit von Höreindrücken und damit auf die Frage nach der Gültigkeit von Hörurteilen zu wissenschaftlichen Zwecken gestoßen (Anders 2001a, Neuber 2002, Kranich 2003, Bose 2001 & 2003). Dieses vermeintliche Problem umgehen Arbeiten, die sich mit rein messphonetischen Analysen dem Sprechausdruck nähern (Kienast 2002, Paeschke, 2003, Schröder 2004) und mit ihren Verfahren statistisch „harte“ Daten ermitteln. Demgegenüber liefern auditive phonetische Analysen eher statistisch „weiche“ Daten, die durch entsprechende methodische Designs (Interview, Fragebogen) in auswertbare Kategorien untergliedert werden müssen (Merkmalsinventar zur Beschreibung von Sprechausdrucksmerkmalen s.o.). Dennoch bezeichnet Bose (2001:269) diese Form der Analyse von Sprechausdrucksmerkmalen als näher an der Realität der Kommunikationspartner liegend und damit auch als zuverlässiger für die Beurteilung musterhafter Sprechausdrucksgestaltung.

Den Ausgangspunkt auditiver Analysen bilden Fähigkeiten und Fertigkeiten, Sprechereignisse mit Hilfe des Gehörs segmental und suprasegmental durch Beurteilung, Skalierung und Transkription in Merkmale und Merkmalskomplexe zu zerlegen (Vieregge 1996:1). Diese Methode ist seit langem etabliert und wird entsprechend als ‚analytisches‘ (Vieregge 1996), ‚kategoriales‘ (Neppert 1999) oder ‚funktionelles‘ Hören (Krech 1960) bezeichnet. Je nach Fragestellung und Untersuchungsziel werden für die auditive Analyse trainierte Hörer (Experten) oder untrainierte Hörer (Laien) herangezogen. In jedem Fall geht es aber darum, anhand eines unverwechselbaren Beurteilungsmaßstabes zu gültigen Hörerurteilen zu kommen, die auf determinierten Erwartungsnormen beruhen.

Bereits frühe Fachvertreter (Dempwolff 1919, Essen 1952:11 & 1979: 237) beschreiben, dass durch die Schulung des Gehörs das phonetische Eichen im Hören möglich sei. Sie beziehen sich dabei auf Beobachtungen, die Zusammenhänge zwischen sprecherischen Merkmalsausprägungen einerseits und der Einschätzung des Sprechausdrucks andererseits aufzeigen. Auch heute noch gilt als Lehrmeinung, dass Fachvertreter auf Grund eines intensiven Hörtrainings in der Lage sind, bei gelenkter und bewusster Aufmerksamkeit, Verlaufs-

formen einzelner Parameter einzuschätzen und damit auf einzelne Sprechausdrucksmerkmale des auditiv wahrnehmbaren Komplexsignals zu achten (Anders 2001a:65, Bose 2001:269). Bose (2001:269) spricht in diesem Zusammenhang auch von:

[...] ‚*bottom-up*‘-Hören im Gegensatz zum normalerweise ganzheitlichen, vorrangig semantisch orientierten ‚*top-down*‘-Wahrnehmungsprozess in der Kommunikation. (Bose 2001:269)

Die Ergebnisse des analytischen Hörens können immer nur einen subjektiven, nie einen objektiven Status beanspruchen, weswegen bei den methodischen Vorüberlegungen und bei der Auswahl der Stichprobe berücksichtigt werden muss, ob es sich um statistisch relevante Größen handelt, die intra- und intersubjektive Übereinstimmungen der Deskription gewährleisten.

Sprechausdrucksparameter wie Stimmhöhe, Intensität, Sprechgeschwindigkeit, Artikulation und ihre jeweiligen Veränderungen scheinen für geschulte Ohren bei der Einschätzung durch eine auditive Analyse keine Schwierigkeiten darzustellen. Kontrovers diskutiert wird allerdings die Frage nach konstanten Hörmustern bei der Beurteilung von Merkmalen des Stimmklangs. Kranich spricht im Zusammenhang mit Timbreeigenschaften von

[...] subjektive[n] Anmutungen, die somit nicht objektivierbar sind. (Kranich 2003:93)

Die Kategorien zur Beschreibung von Stimmklangmerkmalen werden als Metaphern Inventaren anderer Wahrnehmungsbereiche entnommen (s.o.). So findet sich in der Literatur eine breite Palette an Beschreibungsmerkmalen (Luchsinger & Arnold 1970, Wendler & Seidner 1977, Pfau & Streubel 1982, Wirth 1995, Friedrich & Bigenzahn 1995).

Apparativ-phonetische Untersuchungen zu Timbreeigenschaften können zwar noch keine detaillierten Daten liefern, die konkrete Aussagen für einzelne Stimmqualitätsmerkmale zulassen, aber sie können Korrelationen zwischen dem Kriterium der Stimmqualität und spektrografischen Charakteristika aufzeigen (Neppert 1999:65f).

Die Wahrnehmung der Klangfarbe einer Stimme ist oft von der Tonhöhe beeinflusst (höhere Stimmen wirken meist heller), weswegen bei funktionalem Hören, wie es für phonetische Analysen unerlässlich ist, die Tonhöhe und die Helligkeit strikt getrennt werden müssen (Neppert 1999:67).

Unter Gesichtspunkten des funktionalen Hörens kann davon ausgegangen werden, dass geschulte Ohren, die durch ein Hörtraining geeicht wurden, qualitative Aspekte von Timbreeigenschaften wie Helligkeit oder Schärfe (Zwicker 1982:84f, 148ff) überindividuell als Hörmuster erkennen und zuweisen können. Kranich argumentiert vor diesem Hintergrund, dass, obwohl es sich bei den Timbreeigenschaften um sehr subjektive Kriterien handelt,

[...] Stimmklangcharakteristika auditiv relativ einheitlich wahrgenommen werden, weshalb man von durchaus ähnlichen Wahrnehmungsmustern ausgehen kann. (Kranich 2003:94)

Methodisch ergeben sich, je nach wissenschaftlichem Fokus, verschiedene Formen der auditiven Beschreibung. Die einfachste Methode ist dabei, dass Kontrollhörer (Laien) entscheiden sollen, ob eine bestimmte Sprechweise zu einem bestimmten Sprechausdrucksmuster zu rechnen ist oder nicht. Dabei interessieren gemäß des Verständnisses der Sprechwirkung (s.o.), die die generelle Akzeptanz einer Sprechfassung untersucht (Anders 2001a:11), die kommunikative Adäquatheit und Korrektheit einer Äußerung im Sinne von Natürlichkeit, Stilisierung, Echtheit, Anschaulichkeit und Lebendigkeit (Stock & Suttner 1991:69). Die Form der Beurteilung von Sprechfassungen durch Kontrollhörer in standardisierten Hörtests bezeichnet Bose (2001:270) als holistische Musteridentifikation.

Demgegenüber stehen Einschätzungen von Expertengruppen, die als analytische Merkmalsbeschreibung eine Alternative zu Hörtests darstellen. Dabei werden geschulte Hörer gebeten, graduelle Abstufungen verschiedener Eigenschaftsausprägungen auf einer Skala einzuordnen. Beide Methoden ermöglichen die Ermittlung von generellen Hörmustern, wobei individuelle Wahrnehmungen gestützt und intersubjektiviert werden, wodurch allgemeingültige Aussagen anhand eines konkreten Beurteilungsmaßstabes abgeleitet werden können.

Neben den hier genannten stimmlichen und sprecherischen Merkmalen zur Sprechausdrucksbeschreibung soll im folgenden Kapitel aufgezeigt werden, inwieweit auch textinhärente lautsymbolische Phänomene den Sprechausdruck beeinflussen können.

### 4.3 LAUTSYMBOLISMUS & IKONIZITÄT

Im Gegensatz zu den oben beschriebenen indexikalischen Funktionen, die auf realen Beziehungen stimmlicher und sprecherischer Realisationen zur Charakterisierung des Sprechausdrucks beruhen und damit Zeichen beschreiben, die auf Erfahrungen basieren, beschreiben Symbole Beziehungen, die auf Konventionen und Ikone Beziehungen, die auf Ähnlichkeiten beruhen.

Die Frage nach Lautsymbolen und nach der Ikonizität von stimmlichen Phänomenen ist schon in frühen Untersuchungen Gegenstand wissenschaftlichen Interesses gewesen. Dabei wurden vor allem Sprachen als Untersuchungsgegenstand herangezogen, die gängigen Fremdsprachen entfernt sind wie verschiedene afrikanische Sprachen, Chinesisch, Japanisch, Tschechisch (Tsuru/Fries 1933, Brown/ Black/Horowitz 1955, Childs 1994), weil dies die optimale Gelegenheit bietet, bei Versuchspersonen unbewusste Kenntnisse über Lautqualitäten zu ermitteln.

Lautsymbolismen lassen sich in verschiedene Funktionsbereiche typologisieren, wobei Hinton et al. (1994:2f) vier Kategorien benennen, die sie (a) als körperlichen Lautsymbolismus, (b) als imitativen Lautsymbolismus, (c) als synästhetischen Lautsymbolismus und (d) als konventionellen Lautsymbolismus bezeichnen. Unter körperlichem Lautsymbolismus subsumieren die Autoren Ausdrücke emotionaler oder physischer Zustände, die durch die Verwendung von bestimmten Lauten oder Intonationsmustern hervorgerufen werden wie z.B. Husten & Schluckauf, expressive Intonationsverläufe (Paeschke 2003, Kranich 2003, Murray & Arnott 1993, Tolkmitt & Scherer 1986), expressive Stimmqualitätsmerkmale (Kienast 2002, Laver 1980) und die Verwendung bestimmter Interjektionen (Schröder 2000 & 2004). Diese Symptome beziehen sich auf die Inhalte und Studien zur affektiven Prosodie, die in Kapitel 2.2.3. bereits unter dem Aspekt ‚Emotion und Prosodie‘ ausführlich beschrieben wurden.

Der imitative Lautsymbolismus, der auch als Laut- oder Klangmalerei bezeichnet wird, untersucht onomatopoetische Wörter und Ausdrücke, die Geräusche, Laute und Klänge nachahmen bzw. nachbilden (Meid 1999:290f). Dies bezieht sich häufig auf die Darstellung von Tierlauten oder anderer Geräusche, bei denen der Versuch einer akustischen Wiedergabe, z.B. durch die Reduplikation von Silben wie bei ‚wau wau‘ oder ‚kuck kuck‘, vorgenommen wird. Oder es

handelt sich um den Versuch eines akustischen Eindrucks, wie z.B. in quieken, quieksen oder quietschen, bei dem grammatikalische Regeln beibehalten werden. Die Form des imitativen Lautsymbolismus führt Etzel (1983:26) im Zusammenhang mit spontanen Wortschöpfungen an, bei denen es sich um noch nicht domestizierte Wörter handelt, die von einer Sprachgemeinschaft erst akzeptiert werden müssen.

So entwickelt sich in der Ewe-Sprache etwa aus dem Lautbild ‚dabodabo‘ ‚wackligen, watscheligen Ganges‘ ein gleichlautendes (aber in der Tonhöhe verändertes) Nomen ‚Ente‘. Hier drückt die Reduplikation der einfachen Form ‚dabo‘ an sich schon Wiederholung aus, hinzu kommt die Alternation zwischen zwei benachbarten Kontaktpunkten, Zähnen und Lippen, was einem Hin & Her gleichkommt, wie es der bezeichneten Gangart eignet. (Etzel 1983:26)

Eine andere Kategorie des Lautsymbolismus bezeichnen Hinton et al. (1994) als synästhetischen Lautsymbolismus. Hierbei handelt es sich um die akustische Symbolisierung nicht-akustischer Phänomene durch die systematische Verwendung bestimmter Vokale, Konsonanten und prosodischer Eigenschaften zur Darstellung von Eigenschaften wie Größe, Form oder Charakter. So werden beispielsweise durch hohe Vokale und palatale Konsonanten eher kleine Objekte (Diminutiva) und durch dunkle Vokale und Konsonanten eher große und schwere Objekte bezeichnet. Nach Neppert (1999:67) lässt sich die Klangfarbe (hell, dunkel) physikalisch weitgehend der im Schallspektrum begründeten Schallqualität der Sprechlaute zuordnen. Auch die Dauer und tiefe Melodieführung sind Indizien für eher große Objekte. Diesen Aspekt beschreibt Ohala (1983, 1984, 1994) in verschiedenen Arbeiten (vgl. Kap. 3.3.1) indem er die Existenz eines so genannten „frequency codes“ annimmt, wonach scheinbar ungleiche Phänomene eine tiefer liegende Beziehung eingehen. Diese Überlegungen macht er an Parametern wie z.B.  $F_0$  fest, denen sich seiner Meinung nach lautsymbolische Entsprechungen zuordnen lassen. So werden tiefe  $F_0$  Werte mit körperlich großen Individuen und raue Stimmen mit Eigenschaften wie Stärke, Größe und Selbstzufriedenheit verbunden.

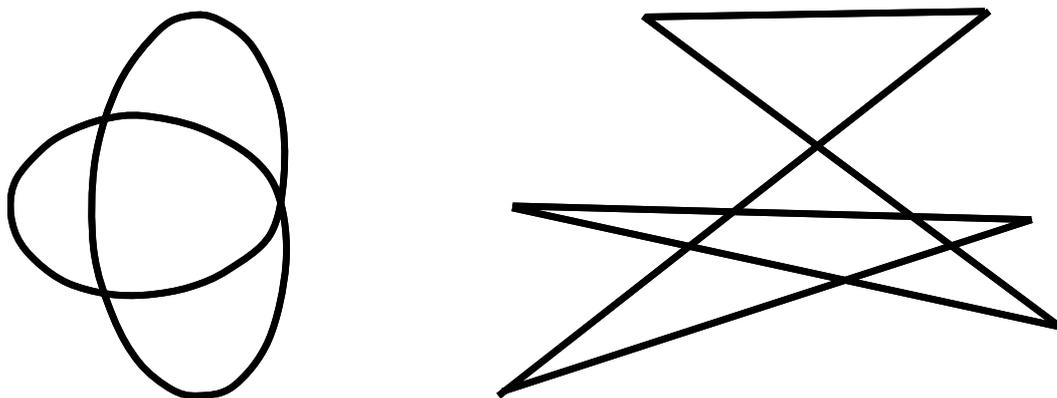
The sound-meaning correlations found in these cases adhere to the ‚frequency code,‘ which also governs the vocalisations of other species, namely, where high  $F_0$  signifies (broadly) smallness, non-threatening attitude, desire for the goodwill of the receiver, ect., and low  $F_0$  conveys largeness, threat, self-confidence and self-sufficiency. (Ohala 1994:343)

Ein frühes Experiment, was die Assoziation von Formen durch Laute und Lautverbindungen klassisch beschreibt, führte Köhler (1929) durch. Dabei legte er

Versuchspersonen die zwei sinnfreien Wörter „maluma“ und „takete“ vor, die sie, zwei ebenfalls sinnlosen Strichfiguren zuordnen sollten (vgl. Abb.9). Aus der Sicht des Beliebigkeitspostulats wäre eine annähernde Gleichverteilung in der Zuordnung zu erwarten.

In den Versuchen, die in verschiedenen Kulturkreisen durchgeführt wurden, ordnete aber die überwiegende Mehrheit der runden Figur (Abb. 9) „maluma“ und der eckigen Figur (Abb. 9) „takete“ zu. Etzel schreibt dazu:

Und in der Tat scheint es spontan plausibel, etwas Rundes (vielfach mit Weichheit assoziiert) mit stimmhaften Nicht-Plosiva zu verbinden, etwas Spitzes, Eckiges dagegen (vielfach mit Härte assoziiert) mit stimmlosen Plosiva zu verbinden. (Etzel 1983:27)



**Abb. 9: Strichfiguren als Grundlage für Lautsymbole (nach Etzel 1983:27)**

Dieses Experiment deutet auf eine Synästhesie hin, die zwischen den akustischen und optischen Gestalten zu bestehen scheint. Und da sich dieses Phänomen über die Kulturkreise hinweg beobachten lässt, scheint es sich bei dieser Analogie um ein allgemeinemenschliches Vermögen zu handeln.

In der sich anschließenden eigenen empirischen Untersuchung spielen solche synästhetischen Lautsymbole ebenfalls eine Rolle. Für ein Männerparfum wurden ein Markenname und ein Slogan kreiert, die sich durch die Prominenz des Vokals /a/ und durch die Alternation des bilabialen bzw. alveolaren Nasalkonsonanten /m/ und /n/ auszeichnen („Caractère. Macht am Mann das Markante am Mann.“). Welche Symbolkraft den einzelnen Lauten bzw.

Lautverbindungen dabei anhaften kann, wird an entsprechender Stelle des Methodenteils (Kap. 5.2) genauer beschrieben werden. In jedem Fall entsteht durch die Konzentration auf eine Vokalqualität in Kombination mit den prominent vertretenen Nasalkonsonanten ein phonetischer Wohlklang, der nur an zwei Stellen durch einen velaren Plosiv aufgebrochen wird, wodurch mit großer Sicherheit der perzeptive Eindruck des Werbespots unter auditiven Gesichtspunkten nachhaltig beeinflusst wird.

Ein anderer Aspekt, der aus Köhlers Ergebnissen resultiert, ergibt sich aus der Tatsache, dass die beschriebenen synästhetischen Effekte sich musterhaft und überindividuell auffinden lassen, was gegebenenfalls auch Rückschluss auf bestimmte Hörmuster zulässt.

Unter der letzten genannten Form des Lautsymbolismus, dem konventionellen Lautsymbolismus, werden gewisse Lautverbindungen oder Sequenzen verstanden, die sprachspezifisch (konventionell) mit bestimmten Bedeutungen assoziiert werden, wie z.B. /gl-/ in glänzen, glitzern oder glühen. Dabei versucht ein Teilgebiet der Psychologie, die Psychophonetik, die mit Sprachlauten verknüpften Empfindungen zu messen und mittels statistischer Untersuchungen einer hinreichend großen Zahl von Probanden die Sprachlaute auf einer Eigenschaftsskala bewerten zu lassen. Gemäß der Methode des semantischen Differentials (Osgood 1952 & Osgood et al. 1957) hat Ertel (1969) alle Einzel-laute in ein zweidimensionales System mit den Koordinaten ( $\pm$ ) Erregung und ( $\pm$ ) Potenz eingeordnet. Im Ergebnis fand er heraus, dass neben dem Einzel-laut auch die Wortform lautsymbolisch beeinflusst ist. Für die Beurteilung des Lautcharakters vermutet er dabei sowohl kinästhetische als auch sensorische Eindrücke. Zur Erklärung der Lautsymbolik stellt sich aber die Frage, ob eher kinästhetische oder eher auditive Momente bei der Lautcharakterbeurteilung von Bedeutung sind. Synthetisierend formuliert Etzel:

Kinästhesie (die Bewegungsempfindung) ist ein sprecherabhängiges Moment, gehört also zur Ausdrucksseite des phonetischen Geschehens, während die Hörempfindung eine Eindrucksqualität ist. Platons Zugang der Deutung der Lautcharaktere ging von der Lautgebärde, dem artikulatorischen Ausdruck aus, wohingegen Ertels psychophonetische Untersuchungen die Eindrucksqualität der Laute zum Gegenstand haben, die im Übrigen nicht nur für den Hörer, sondern auch für den sich selbst hörenden Sprecher von Einfluß sind. (Etzel 1983:36f)

Das Abtragen von Eigenschaften in Form binär angeordneter Paare bietet somit neben der Möglichkeit der Klassifizierung auch eine solide Basis für die Darstellung von Wertigkeiten.

#### 4.4 ZUSAMMENFASSUNG

Sprecherische Ausdrucksformen, die einerseits individuumsspezifisch ausgeführt werden, unterliegen andererseits in ihrer Gestaltung situations- und stimmungsadäquaten Konventionen, was in der Sprechwissenschaft (Krech et al. 1991) unter dem Begriff ‚Sprechausdruck‘ zusammengefasst wird. Dabei lassen sich verschiedene Betrachtungsweisen unterscheiden. Unter der externen Rezeptionsdeterminante des Sprechausdrucks werden sowohl indexikalische als auch lautsymbolischen Phänomene verstanden, die anhand bestimmter Zeichen, Symbole oder Ähnlichkeiten Auskunft über pathognomische und physiognomische Gestaltungsweisen einer Äußerung geben. Dies geschieht durch bewusstes oder unbewusstes Einsetzen und Variieren sowohl segmentaler als auch suprasegmentaler Merkmale, die sich in Form eines Kriterienkataloges zusammenstellen lassen.

Bei der Analyse von Sprechausdrucksmerkmalen wird versucht, Muster aufzuzeigen, die sich sprecher- und hörerüberindividuell auffinden lassen. Die interne Rezeptionsdeterminante, die dabei versucht, Hörmuster anhand von Hörgewohnheiten herauszukristallisieren, nähert sich dieser Problematik methodisch durch funktionelles Hören von untrainierten Hörern (Laien) oder trainierten Hörern (Experten) an. Verschiedene Arbeiten (Anders 2001a, Bose 2001 & 2003, Kranich 2003) konnten nachweisen, dass diese Methode, die immer nur einen subjektiven Status beanspruchen kann, zuverlässige Ergebnisse in Bezug auf die Existenz intra- und intersubjektiver Übereinstimmungen von Hörgewohnheiten liefert.

Sowohl die Analyse externer als auch die Untersuchung interner Rezeptionsdeterminanten stellen wichtige phonetische Methoden dar, um Sprechausdrucksphänomene zuverlässig und musterhaft zu beschreiben. Im folgenden Teil der Arbeit wird dies an einer eigenen Untersuchung zur Perzeption emotionalisierter Sprechweisen am Beispiel einer Parfumwerbung aufgezeigt werden.

## 5 EXPERIMENTELLES DESIGN

### 5.1 FRAGESTELLUNG UND ZIELSETZUNG

In diesem Teil der Arbeit soll eine eigene Untersuchung vorgestellt werden, die der Frage nach der Perzeption prosodischer Merkmale im Persuasionsprozess einer Parfumwerbung nachgeht. Ziel der Untersuchung ist es herauszufinden, ob die Veränderung stimmlicher und sprecherischer Merkmale verschiedener Sprechweisen eines Werbeslogans zu einer veränderten Wahrnehmung und damit Wirkung eines Werbespots führt. Dass die Stimme als Wirkungsfaktor in den Perzeptionsprozess einer Äußerung involviert ist, belegen verschiedene Arbeiten (z.B. Eckert & Laver 1994, Gélinas-Chebat et al. 1996, Gélinas-Chebat & Chebat 1999), aber inwieweit prosodische Veränderungen zu einer Änderung der Wahrnehmung führen, ist in der Literatur hauptsächlich im Zusammenhang mit der Diskriminierung einzelner Emotionen beschrieben (z.B. Kienast 2002, Paeschke 2003, Kranich 2003). Die Identifikation und Zuweisung von Emotionen anhand prosodischer Bestandteile einer Äußerung bildet im Zusammenhang mit der hier vorliegenden Untersuchung allerdings keinen Schwerpunkt, vielmehr soll transparent gemacht werden, inwieweit markierte Sprechweisen Einfluss auf bestimmte Wahrnehmungsbereiche nehmen können. Das bedeutet im konkreten Fall die Frage danach, ob es möglich ist, durch eine natürliche Manipulation der Stimme und Sprechweise auch Veränderungen in der Wahrnehmung des gesamten Werbespots, in der Wahrnehmung der dargestellten Person oder in der Wahrnehmung des assoziierten Duftes auszumachen.

Die gewählte Methodik wird im folgenden Kapitel näher beschrieben. Geleitet wurde das Design der Untersuchung von nachstehenden Fragestellungen:

- Welchen Einfluss haben ausgewählte Parameter der Stimme wie Sprechstimmhöhe, Intensität, Klangfülle, Klangfarbe und Stimmqualität auf die Beurteilung einer Parfumwerbung?
- Welchen Einfluss hat die Artikulation auf die Wahrnehmung und Beurteilung eines Werbevideos?
- Welchen Einfluss haben einzelne Verlaufsqualitäten der Sprechweise wie Tonhöhenschwankungen ( $F_0$ -Verlauf), Sprechgeschwindigkeit, Dehnungsakzente und Pausen auf die Beurteilung eines Werbespots?

- Gibt es Parameter in der Stimme und Sprechweise, die eine Veränderung der Wahrnehmung des Werbevideos verstärkt beeinflussen?
- Verändert sich durch die Modulation der Stimme und der Sprechweise die Wahrnehmung des gesamten Werbespots, oder werden insbesondere einzelne Elemente (z.B. dargestellte Person, assoziierter Duft) verändert wahrgenommen?

Ziel der Arbeit ist es also, herauszufinden, ob und inwieweit Veränderungen der Stimme und Sprechweise in einem Werbespot zu einer veränderten Perception (Wahrnehmung des Werbespots) und Persuasion (Einstellung zum Produkt/Einstellung zum Kauf des Produkts) führen. Aus dieser Zielstellung ergibt sich die folgende Hypothese, die der Untersuchung vorangestellt werden soll.

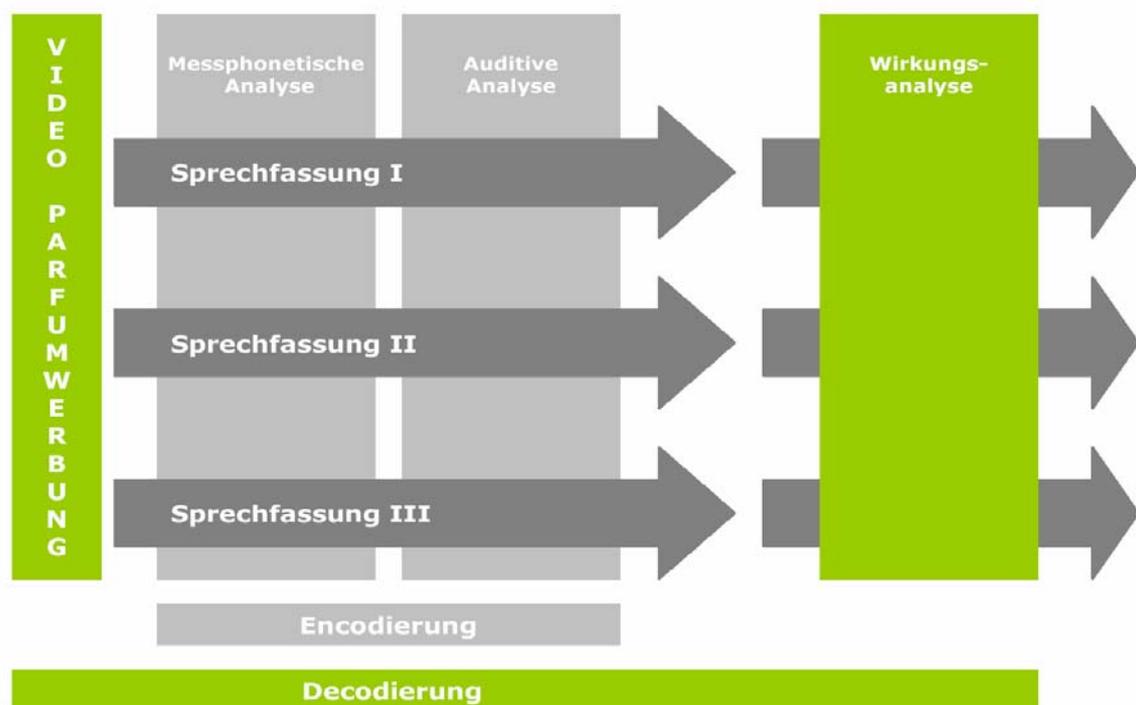
**HYPOTHESE:**

H<sub>0</sub>: Veränderungen von stimmlichen Parametern und sprecherischen Realisationen eines Werbeslogans beeinflussen die Perception und die Persuasion eines Werbespots nicht.

H<sub>1</sub>: Veränderungen von stimmlichen Parametern und sprecherischen Realisationen eines Werbeslogans beeinflussen die Perception und die Persuasion eines Werbespots.

## 5.2 MATERIAL UND METHODEN

Um den in Kapitel 5.1 genannten Fragen nachzugehen, wurde eine Synthese verschiedener methodischer Ansätze gewählt, die sowohl messtechnisch als auch auditiv ermittelte Daten aufgreift und in ein komplexes Modell integriert (Abb.10). Dies geschieht in Anlehnung an Scherers Adaption des Brunswik-schen Modells (vgl. Kap. 3.2), und lässt sich folgendermaßen auf die hier vorliegende Fragestellung übertragen:



**Abb. 10: Schematische Darstellung des experimentellen Designs in Anlehnung an Scherer (1978).**

### Werbevideo

Den Ausgangspunkt der Untersuchung bildete ein bearbeitetes Werbevideo der Marke „Giorgio Armani“ des Männerparfums „ACQUA DI GIÒ“, welches drei separaten Probandengruppen jeweils drei Mal hintereinander vorgespielt wurde, wobei sich die gezeigten Videos ausschließlich in der Sprechweise unterschieden. Die Probanden sollten das jeweilige Video im Anschluss an die Prä-

sentation auf einem Fragebogen anhand verschiedener Kriterien einschätzen (s.u. und siehe Anhang).

Das Originalvideo wurde zu diesem Zweck unkenntlich gemacht, indem der Name des Produkts, der Slogan, der Flakon und die Sprechweisen so verändert wurden, dass scheinbar ein neues Parfum entstand, welches in Wirklichkeit auf dem Markt nicht existiert und nicht beworben wird. Dieser Entscheidung lag die Überlegung zu Grunde, dass die Probanden ein Video beurteilen sollten, welches sie aus dem Fernsehen noch nicht kennen konnten, um sicher zu gehen, dass keine bereits vorhandenen Einstellungen (Sympathien oder Antipathien) bezüglich des Parfums abgefragt würden, sondern dass die Probanden ausschließlich das präsentierte Videomaterial beurteilten (vgl. ELM in Kap. 2.2). Die Bearbeitung des Videos erfolgte durch einen Mediendesigner unter Verwendung verschiedener Bildverarbeitungsprogramme der Marke Adobe<sup>®</sup>, was den Vorteil der Kompatibilität der Dateiformate zwischen den verschiedenen Tools mit sich brachte. Für die Bearbeitung des Bildmaterials wurden After effects<sup>®</sup> für das Compositing (Farbkorrekturen, Einblenden der Schrift, Rauschfilter, Blenden setzen), Photoshop<sup>®</sup> für die Bildbearbeitung (Flakon gesäubert und neu beschriftet) und Premiere<sup>®</sup> für das Codieren von verschiedenen Videoformaten verwendet. Die Nachbearbeitung der Musik und des Tons erfolgte mit Steinberg wavelab<sup>®</sup>, wobei die Musik geloopt und die separat aufgenommenen Sprechfassungen entsprechend eingebaut wurden. Um abzusichern, dass die Probanden das Originalvideo, falls bekannt, nicht assoziierten, wurde dies ebenfalls im Fragebogen erhoben, wobei 90% aller Teilnehmer<sup>9</sup> angaben, dass Parfum und die Werbung nicht zu kennen.

Der neu kreierte Werbespot umfasste bei einer Gesamtdauer von 20 Sekunden fünf Bildsequenzen (Abb. 11), die inhaltlich fließend ineinander übergingen und setzte sich aus folgender Bildfolge zusammen: Aus den Strudeln des bewegten Meeres (Motiv 1) stieg ein dynamischer männlicher Körper hervor (Motiv 2), dessen große und elegante Armbewegungen beim Delphinschwimmen die Bewegungen des Meeres gleichsam aufgriffen und dadurch besonders athletisch wirkten. Diesen Bewegungsbildern folgte über eine Blende ein Schnitt zu einer Portraitaufnahme des Gesichtes des Protagonisten (Motiv 3), welches durch eine androgyne, knabenhaft-zeitlose Schönheit charakterisiert war und über welches einzelne Wassertropfen perlten. Eingebildet aus dem Anblick

---

<sup>9</sup> Für genauere Angaben zur Beschreibung der Stichprobe siehe auch Kapitel 6.3.

des Gesichtes trat für drei Sekunden ein eleganter, formstrenger und unverspielter Flacon des Parfums „Caractère“ hervor (Motiv 4), der bereits nach 2 Sekunden wieder im Nichts verschwand. Zurück blieb eine schwarze Fläche (Motiv 5).

Erst bei Motiv 5 erklang eine männliche Stimme, die vor dem schwarzen Hintergrund den ellipsenhaften, stark minimalistischen Slogan „Caractère. Macht am Mann das Markante am Mann“ artikulierte, wobei der Name des Parfums „Caractère“ und „das Markante am Mann“ in schlanker weißer Schrift vor dem schwarzen Hintergrund erschienen. Die Sprechweise variierte je nach Sprechfassung zwischen einer Länge von 3.35 Sekunden und 4.19 Sekunden, und nahm dabei mit ca. einem Fünftel der Gesamtdauer des Werbespots, nur einen sehr kurzen Präsentationsraum innerhalb des gesamten Werbespots ein.

Durch einen strengen Minimalismus, vor allem in der Verwendung weniger, aber sehr exklusiver Bilder und Motive wurde versucht, einen Gedanken von Elite und Exklusivität zu transportieren, der gleichzeitig signalisieren sollte, dass es sich bei diesem Produkt um ein sehr hochwertiges Parfum handele, welches einer bestimmten Zielgruppe, aber keiner großen Allgemeinheit zugesprochen werden sollte.

Minimalistisch wirkte der Werbespot an erster Stelle durch die Wahl der Darstellung in schwarz-weiß-Bildern, welche über die gesamte Zeit durch dezente wellenförmige und dynamische Melodiebewegungen untermalt wurden. Diese waren wiederum durch sehr kleine Tonhöhenintervalle und ausgeprägte rhythmische Akzente charakterisiert, so dass man einzelne Motive ständig wieder zu erkennen glaubte, die sehr gleichförmig anmuteten, aber dennoch unterschiedlich waren. Genau durch diese scheinbar vorhandenen musikalischen Schleifen wurde der sehr bewegte Charakter des Meeres musikalisch nachempfunden und assoziiert.

Bild, Musik und Sprechweise des Werbespots waren fließend miteinander verbunden und ließen sich in ihrer Wirkung auch nicht voneinander trennen. Über die Bedeutung von Bild und Musik in der Werbung existieren verschiedene Arbeiten (z.B. Schierl 2004 & 2001, Kloß 2000, Wüsthoff 1999, Kroeber-Riel 1996, Tauchnitz 1990, Helms 1981). In dieser Arbeit wird zur Kenntnis genommen, dass alle Faktoren maßgeblich zusammenwirken, die Konzentration liegt fortan allerdings ausschließlich auf der Untersuchung von stimmlichen

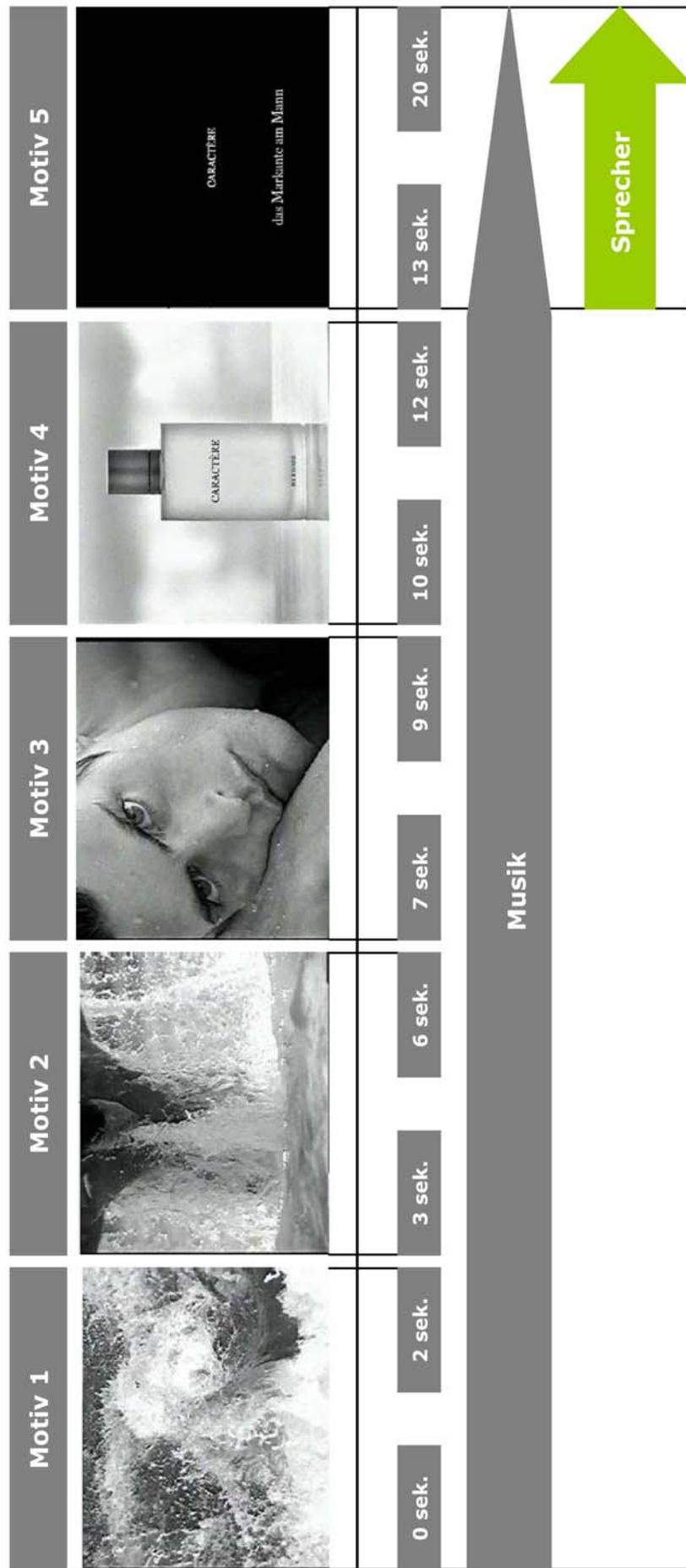


Abb. 11: Bildfolgen und Filmsequenzen des Werbespots.

und sprecherischen Einflüssen bezüglich der Wahrnehmung einer Werbung. Würde man aber beispielsweise die Sprechweisen isoliert präsentieren und einschätzen lassen, um andere Faktoren wie Bild und Musik in ihrer Wirkung auszuschließen, so würde sich die Beurteilung dahingehend verändern, dass die Werbung als nicht authentisch und unecht beurteilt würde. Von daher wurden drei verschiedene Sprechfassungen und damit Videos erstellt, die sich ausschließlich in der Stimmqualität und Sprechweise voneinander unterscheiden und jeweils drei separaten Gruppen vorgespielt, so dass Veränderungen in der Beurteilung des Werbespots allein auf Veränderungen der Stimme und Sprechweise zurückgeführt werden konnten, da alle anderen Parameter streng konstant gehalten wurden.

Abbildung 11 zeigt eine schematische Darstellung der einzelnen Bestandteile des untersuchten Werbespots und eine Abfolge der Bildsequenzen und Motive. Zusätzlich sind auf einem Zeitstrahl die Dauern der einzelnen Motive in Sekunden abgetragen, was die Proportionen zwischen Bild, Musik und Anteil der Stimme transparenter gestaltet.

Für die Sprechfassungen wurde ein erfahrener und professioneller männlicher Sprecher gewählt, der über ein differenziertes Repertoire an Modulationsfähigkeiten der Stimme und Sprechweise verfügt. Er bekam die Anweisung, den Slogan „Caractère. Macht am Mann das Markante am Mann.“ in drei Sprechweisen zu sprechen, die sich deutlich hörbar voneinander unterscheiden. Gelenkt wurde die Aufmerksamkeit dabei auf die Vorstellung verschiedener Düfte. Über den Versuch einer synästhetischen Wirkung durch assoziative Nähe, sollten drei Sprechfassungen entstehen, die sich olfaktorisch als (a) frisch und leicht, (b) athletisch und dynamisch und (c) würzig und schwer anhörten. Für die genaue sprecherische Umsetzung wurden keine konkreten Vorgaben gemacht. Die gesprochenen Versionen wurden mit CSL, der Firma Kay Elemetrics® als wav-Dateien bei 48 kHz und 16 BIT unter Laborbedingungen aufgezeichnet.

Bei der Wahl des Namens des Parfums und des Slogans wurde darauf geachtet, dass ein phonetischer Wohlklang entstand, der sich gut in den Gesamtkontext des Werbespots integrieren ließ. Auffällig und beabsichtigt war die Prominenz des Vokals /a/ und die Akkumulation der Nasalkonsonanten /m/ und /n/. Der daraus resultierende weiche und fließende Charakter wurde lediglich an drei Stellen durch den velaren Plosiv /k/ unterbrochen, was gleichsam

das Markante darstellte und damit das Parfum vor seinem Hintergrund kontrastierte. Dies ist in Abbildung 12 noch einmal schematisch dargestellt.

---

**Caractère. Macht am Mann das Markante am Mann.**

---

■ Vokal /a/      ■ Nasale /m/ und /n/      ■ Plosiv /k/

---

**Abb. 12: Lautsymbolik des Namens des Parfums und des Slogans.**

### Messphonetische und auditive Analyse

Die drei aufgezeichneten Sprechfassungen wurden im Anschluss einer messphonetischen Analyse (vgl. Kap. 6.1) und einer auditiven Analyse (vgl. Kap. 6.2) unterzogen, um auf diese Weise die Unterschiede in der Stimmqualität und in den sprecherischen Realisierungen der einzelnen Fassungen herauszuarbeiten. Diese Analysen dienten gleichsam als Zwischenschritte in der Encodierungsphase der Wirkungsanalyse. Das Hauptaugenmerk lag am Ende auf der Fragebogenuntersuchung (Dekodierungsphase), die sich im weiteren Verlauf des Experimentes anschloss (Kap. 6.3), wobei die Probanden vorab in keiner Weise für die Beobachtung der Stimme und Sprechweise sensibilisiert wurden (Blindstudie).

In der apparativen Analyse wurden verschiedene Parameter messtechnisch erhoben, wie:  $F_0$ -Verlauf, Lautdauern, Sprechgeschwindigkeit, artikulatorische Spezifika und Pausen. Diese Parameter wurden mit Hilfe der Programme Wavesurfer<sup>®</sup> und Praat<sup>®</sup> extrahiert und analysiert. In Bezug auf die Segmentierung der einzelnen Laute wurde großer Wert auf Konsistenz bei der Etikettierung gelegt (Simpson 1998:29ff). Geleitet war die Etikettierung dabei von audio-visuellen Kriterien. In schwierigen Fällen (z.B. beim Übergang zwischen dem alveolaren Frikativ und dem bilabialen Nasal in „das **M**arkante am Mann“) wurde als Grenze unter dem Gesichtspunkt der Beständigkeit immer die einsetzende Stimmhaftigkeit angenommen (Kap. 6.1).

Da sich aber nicht alle Stimmeigenschaften (z.B. das Timbre) messtechnisch ermitteln ließen und, da die Messwerte nicht immer mit der menschlichen Wahrnehmung korrelierten, wurden die Sprechfassungen zudem einer auditiven Analyse unterzogen. Um das menschliche Ohr als Analysator verwenden

zu können, bedurfte es eines gründlichen Hörtrainings zur Eichung der unterschiedlichen Hörgewohnheiten, die geprägt durch einen holistischen Verarbeitungsmodus zunächst die Trennung sonst stets kombinierter suprasegmentaler Eigenschaften überwinden lernen mussten (Kranich 2003:72f). Diese Methode der Beurteilung von Sprechfassungen kann nur als verbindlich angeführt werden, wenn die auditive Einschätzung durch Experten geschieht, deren Professionalität in der Beurteilung von Stimmen und Sprechleistungen liegt, wie dies etwa bei ausgebildeten und praktizierenden Sprechwissenschaftlern der Fall ist. Die hier erstellten Sprechfassungen wurden von 6 aktiven Sprecherziehern in Form eines Fragebogens (siehe Anhang) eingeschätzt. Ziel der Analyse war es, Kriterien der Stimme und Sprechweise zu beurteilen, die sich in den drei Sprechfassungen unterschieden, um auf diese Weise übersubjektive Urteile bezüglich der Verschiedenheit in der Wahrnehmung der einzelnen Sprechfassungen zu erhalten. In Anlehnung an Fährmann (1967) wurde dazu zwischen (a) habituellen Stimmqualitäten, die sich auf die scheinbar physiologischen Vorraussetzungen des Sprechers und (b) individuellen Verlaufsqualitäten, die sich auf die sprecherische Realisierung der Äußerung beziehen, unterschieden. Der verwendete Fragebogen (siehe Anhang) setzte sich aus Items zusammen, die von Fährmann (1967) und Kranich (2003) in geeigneter Auswahl für die Fragestellung der eigenen Untersuchung übernommen wurden (Tab. 8).

<b>HABITUELLE STIMMQUALITÄTEN:</b>	<ul style="list-style-type: none"> <li>▪ <b>Sprechstimmlage:</b> tief, mittel, hoch</li> <li>▪ <b>Sprechgeschwindigkeit:</b> langsam, mittel, schnell</li> <li>▪ <b>Intensität:</b> gering, mittel, hoch</li> <li>▪ <b>Stimmqualität:</b> klar, verhaucht, knarrend</li> <li>▪ <b>Klangfülle:</b> resonanzreich, mittel, resonanzarm</li> <li>▪ <b>Klangfarbe:</b> a) hell, mittel, dunkel, b) weich, mittel, hart, c) warm, mittel, kalt</li> <li>▪ <b>Artikulatorische Spezifika</b></li> </ul>
<b>INDIVIDUELLE VERLAUFSQUALITÄTEN:</b>	<ul style="list-style-type: none"> <li>▪ <b>Melodische Veränderlichkeit:</b> quantitativ, qualitativ</li> <li>▪ <b>Dehnungsakzente:</b> quantitativ, qualitativ</li> </ul>

**Tab. 8: Kriterienkatalog für die auditive Analyse des Sprechmaterials in Anlehnung an Fährmann (1967) und Kranich (2003).**

## Wirkungsanalyse

Sowohl die apparative als auch die auditive Analyse dienten als Zwischenschritte der Beschreibung der einzelnen Sprechweisen, um auf diese Weise der Frage nach der Persuasion und Perzeption verschiedener Sprechweisen am Beispiel einer Parfumwerbung nachzugehen. Die für die Wirkungsanalyse benötigten Daten sind in Form eines Fragebogens (siehe Anhang) erhoben worden, wobei die Probanden weder explizit noch implizit für die Beurteilung stimmlicher und sprecherischer Realisationen sensibilisiert wurden. Es handelte sich also um eine Blindstudie, bei der ein systematischer Einfluss des Untersuchungsfokus auf das Verhalten der Versuchspersonen ausgeschlossen wird. Die Aufmerksamkeit der Probanden wurde vielmehr auf Faktoren gelenkt, die offensichtlich mit dem Werbespot und der scheinbar beabsichtigten Wirkung in Verbindung standen. Zu den Kategorien a) Werbespot, b) Person, und c) Duft wurden Attribute in Form von Adjektiven gesucht, die eine mögliche Wirkung beschrieben. Diese Adjektive wurden so gewählt, dass sie sich im Idealfall zu zwei (Werbespot, Person) bzw. drei (Duft) Oberkategorien zusammenfassen ließen, was empirisch durch eine entsprechende Faktorenanalyse überprüft wurde (s.u.). Bei einer Faktorenanalyse werden zugrunde liegende Konstrukte oder Faktoren identifiziert, die die Korrelation zwischen einem Set von Variablen erklären. Demnach gehörten verschiedene Adjektive zu einer übergeordneten Idee und konnten als Bestandteile eines jeweiligen Faktors betrachtet werden. In Tabelle 9 sind die erfragten Items der jeweiligen Kategorien überblicksartig dargestellt, die im Fragebogen durchgängig anhand einer 5-er Skala (von -2 [trifft gar nicht zu] über -1, 0, +1 bis +2 [trifft völlig zu]) beurteilt werden sollten.

KATEGORIE	EIGENSCHAFTEN
(a) Werbespot	exklusiv, edel, farblos, professionell, schlicht, preiswert, animierend, gewöhnlich
(b) Person	charmant, dynamisch, erotisch, weich, aufdringlich, charismatisch, männlich, abstoßend
(c) Duft	blumig, moschusartig, lieblich, markant, süß, tabakartig, dezent, warm, herb, fruchtig, maskulin, leicht, elegant, frisch, holzig, sportlich, würzig, spritzig, klassisch, kühl, vitalisierend, schwer

**Tab. 9: Kategorien des Fragebogens.**

Die Anordnung der Adjektive erfolgte durch ein Zufallsprinzip, so dass sich kein alternierendes System erkennen ließ.

An der Untersuchung nahmen insgesamt 156 Probanden (Studenten und Mitarbeiter der FSU Jena) teil, die sich, wie Tabelle 10 zeigt, in folgenden Zusammensetzungen auf die drei verschiedenen Sprechfassungen aufteilten. Neben diesen Daten wurden noch einige biografische und demografische Daten erhoben, die sich aus Angaben zur Person wie Alter, Geschlecht, Studiengang (Fakultät), Herkunft (Großstadt, Kleinstadt, Dorf), Typ (akustisch, visuell, olfaktorisch) zusammensetzten. Diese Angaben gaben vor allem für die Beschreibung der Stichprobe relevante Aufschlüsse.

<b>SPRECHWEISE</b>	<b>PROBANDEN</b>
(I) frisch & leicht	47
(II) athletisch & dynamisch	75
(III) würzig & schwer	34
<b>Gesamt</b>	<b>156</b>

**Tab. 10: Stichprobe der Fragebogenuntersuchung.**

Die gewonnenen Rohdaten wurden zunächst in SPSS<sup>®</sup> eingearbeitet und verschiedenen Analysen unterzogen. Im Zusammenhang mit der univariaten Varianzanalyse wurden die deskriptiven Statistiken (Mittelwert und Standardabweichung), sowie der Levene-Test zur Prüfung der Konstanz der Fehlervarianzen einer abhängigen Variabel über die Gruppe hinweg, ermittelt. Dem schloss sich ein Post-Hoc-Test (Tamhane und Tukey-HSD) an, der basierend auf beobachteten Mittelwerten Mehrfachvergleiche (Gruppenvergleiche) einer abhängigen Variable ermöglichte. Durch dieses Testverfahren ließen sich Aussagen über signifikante Unterschiede eines Items oder eines Faktors über die verschiedenen Gruppen (Sprechfassungen) hinweg ermitteln.

Darüber hinaus wurde durch einen Chi-Quadrat-Test (Pearson), der zur Überprüfung von Häufigkeitsverteilungen bei Variablen mit nominalem Skalenniveau dient, untersucht, ob sich die Einstellung zum Kaufverhalten über die verschiedenen Sprechfassungen hinweg veränderte, oder inwieweit die Einstellung zum Kaufverhalten auch geschlechtsspezifisch differenziert betrachtet werden musste.

## 6 ERGEBNISSE

### 6.1 APPARATIVE ANALYSE

Für das Video sind drei Sprechfassungen (fortan mit SF abgekürzt) erstellt worden, die sich bei gleich bleibendem Inhalt<sup>10</sup>, hörbar voneinander unterscheiden. Diese entstanden über den Versuch einer synästhetischen Wirkung durch assoziative Nähe und sollten folgende Düfte charakterisieren: (a) frisch und leicht für SF I, (b) athletisch und dynamisch für SF II und (c) würzig und schwer für SF III (vgl. Kap. 5.2).

In diesem Kapitel soll anhand verschiedener phonetischer Analysen aufgezeigt werden, wie sich die sprecherischen Unterschiede im akustischen Signal nachweisen lassen. Das Hauptaugenmerk liegt dabei auf der Beschreibung temporaler, melodischer und artikulatorischer Eigenschaften.

#### **Temporale Struktur und artikulatorische Merkmale**

In allen drei SF wird die inhaltliche Äußerung in drei Sprechsequenzen, die jeweils durch eine Pause voneinander getrennt sind, unterteilt, so dass sich folgendes Schema ergibt: „Caractère [Pause] Macht am Mann [Pause] das Markante am Mann“.

Abbildung 13 a-c zeigt die Gesamtzeit ( $t_{ges}$ ) aller drei SF, die sich aus der Sprechzeit und den jeweiligen Pausen zusammensetzt. Dabei variieren SF I (3.35s) und FS III (4.19s) mit 0.84s Differenz am stärksten. FS II nimmt mit 4.03s die Mittelposition ein. Mit nur 0.16s Unterschied liegen FS II und FS III in ihrer temporalen Struktur scheinbar eng beieinander. Abweichungen zwischen diesen beiden Fassungen zeigen sich aber vor allem, wenn man die Pausenzeiten mit in Betracht zieht.

---

<sup>10</sup> Der Slogan „Caractère. Macht am Mann das Markante am Mann.“ weist auf verbaler Ebene keine persuasive Elemente (Fakten oder Argumente, die für das Parfum sprechen) auf, so dass von der Annahme ausgegangen werden kann, dass die prosodischen Anteile maßgeblich die persuasiven Effekte konstituieren.

**Sprechfassung I**

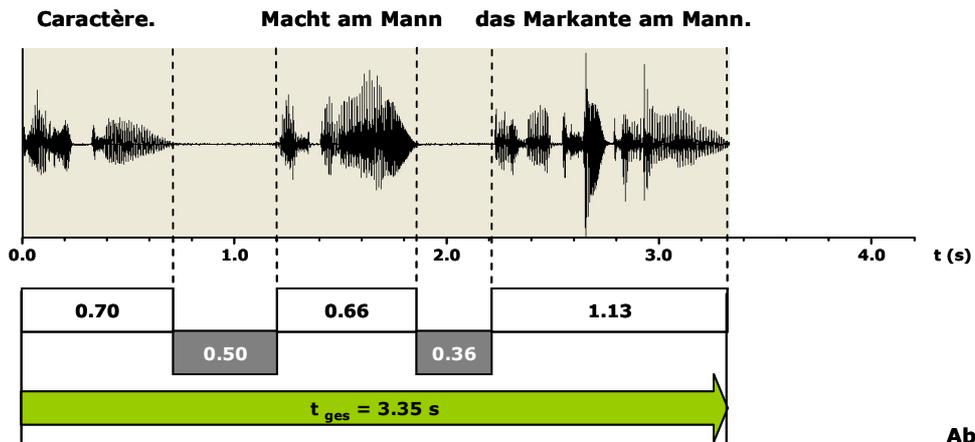


Abb. 13 a

**Sprechfassung II**

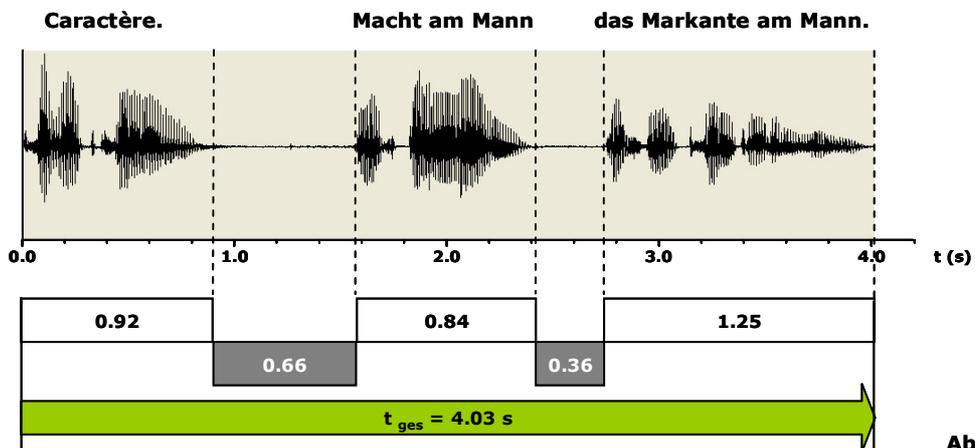


Abb. 13 b

**Sprechfassung III**

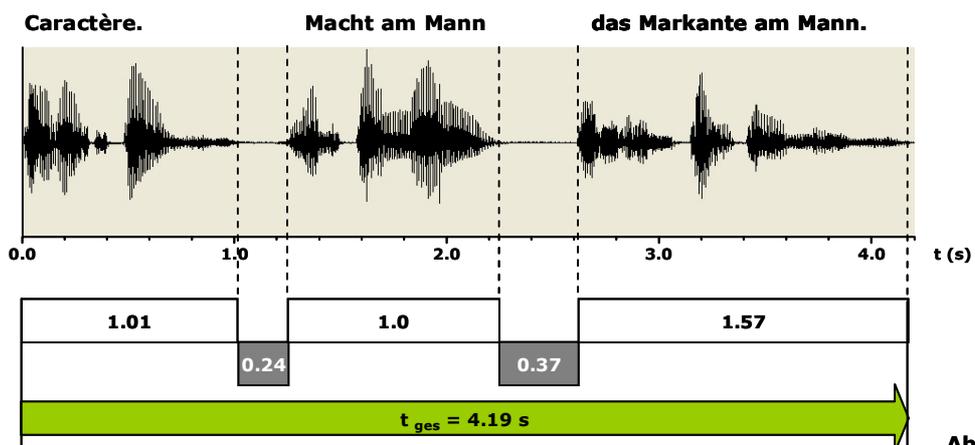


Abb. 13 c

Abb. 13 a-c: Temporale Struktur aller drei Sprechfassungen im Vergleich. Alle Angaben in Sekunden (s). Die grau unterlegten Zahlen kennzeichnen die Pausen.

Über alle drei SF hinweg ergibt sich eine nahezu identische Zeit für die Pause zwischen „[...] Macht am Mann [Pause] das Markante am Mann“, die in SF I und SF II jeweils 0.36s und in SF III 0.37s umfasst (vgl. Abb. 13 a-c). Im Gegensatz dazu weist die erste Pause, die sich zwischen „Caractère. [Pause] Macht am Mann [...]“ befindet, deutlich größere Unterschiede bezüglich ihrer Realisation auf. Mit 0.50s nimmt sie in SF I eine mittlere Position ein, länger wird sie in SF II mit 0.66s und wesentlich kürzer in SF III mit 0.24s realisiert. Kompensiert wird diese kurze Pause in SF III durch eine auffällig starke Dehnung des vorausgehenden Frikativs [χ] in „Caractère“ (vgl. Abb. 13 c).

Die erste Sprechsequenz („Caractère“) und die sich anschließende Pause, die gleichsam den Markennamen und den Slogan voneinander trennen, scheinen maßgeblich für die Wahrnehmung der Sprechgeschwindigkeit aller drei Sprechsequenzen verantwortlich zu sein. So ergeben sich aus der Addition der Sprechzeit von „Caractère“ und der folgenden Pause Zeiten, die für SF I bei 1.20s, für SF II bei 1.58s und für SF III bei 1.25s liegen (Abb. 13 a-c). Diese Rangfolge deckt sich mit den Ergebnissen der auditiven Analyse (vgl. Kap. 6.2), wobei SF III einheitlich als schneller als SF II und als langsamer als SF I beurteilt wurde (vgl. auch Kap. 7).

Vergleicht man nun den gesamten Pausenanteil der einzelnen Realisationen in Bezug zur gesamten Sprechzeit, so ergeben sich rechnerisch durchschnittliche Pausenwerte, die in SF I mit 25.7% und SF II mit 25.3% nahezu identisch sind und sich von SF III mit 14.6% deutlich unterscheiden.

Aufschlussreich bezüglich der Beschreibung der temporalen Struktur erweist sich auch der Vergleich der Sprechanteile innerhalb der einzelnen Sprechsequenzen (Caractère | Macht am Mann | das Markante am Mann). Dabei setzen sich die ersten beiden Sequenzen aus jeweils 3 Silben und die dritte Sequenz aus 6 Silben zusammen. Sowohl fassungsintern als auch über die SF hinweg lassen sich zwei Parallelen aufzeigen: (a) die erste Sprechsequenz (Caractère) nimmt in allen SF in Bezug auf die temporale Beschaffenheit die mittlere Position ein, Sequenz zwei (Macht am Mann) weist eine kürzere und Sequenz 3 (das Markante am Mann) eine längere Realisationszeit auf und (b) die einzelnen Sprechsequenzen wachsen in ihrer temporalen Ausdehnung proportional von SF I über SF II zu SF III, d.h.

jede Sequenz von SF II ist länger als ihre Entsprechung in SF I und kürzer als die Parallelstelle von Fassung III.

Auf dieser Grundlage ergeben sich Sprechgeschwindigkeitswerte für die Sequenzen, die sich in Silben pro Sekunde ermitteln lassen und, die die Variabilität des Sprechtempos transparent machen. Da es sich um sehr kurze Sprechseinheiten handelt, werden mit den Angaben keine durchschnittlichen Sprechgeschwindigkeiten, sondern vielmehr agogische Veränderungen beschrieben. In Tabelle 11 sind die errechneten Werte für die einzelnen Sprechsequenzen und für die gemittelte Gesamtzeit in Silben pro Sekunde dargestellt.

	<b>Caractère</b>	<b>Macht am Mann</b>	<b>Das Markante am Mann</b>	<b>MW <math>t_{ges}</math></b>
<b>Sprechfassung I</b>	4.28	4.55	5.31	4.71
<b>Sprechfassung II</b>	3.26	3.57	4.80	3.88
<b>Sprechfassung III</b>	2.97	3.00	3.83	3.27

**Tab. 11: Sprechgeschwindigkeit innerhalb der einzelnen Sprechsequenzen aller drei Sprechfassungen im Vergleich. In der 4. Spalte ist der Mittelwert der Sprechgeschwindigkeit für alle drei Sprechsequenzen angegeben. Alle Angaben in Silben pro Sekunde.**

Um die drei SF bezüglich ihrer temporalen Eigenschaften hinreichend beschreiben zu können, ist es unerlässlich detaillierter in die einzelnen Sprechsequenzen hineinzuschauen, denn wichtige Diskriminierungsindizes bezüglich der temporalen Beschaffenheit der drei SF finden sich vor allem in fassungsinernen kleineren Einheiten. Dies betrifft beispielsweise die Analyse von Lautdauern, die nicht selten mit artikulatorischen Spezifika wie Glottalisierung, starker Friktion, fehlender Aspiration u.a. einhergehen, und die Beschreibung koartikulatorischer Erscheinungen.

In Abbildung 14 a-c sind die Lautsegmentierungen und deren temporale Ausdehnung für die erste Sprechsequenz („Caractère“) dargestellt. Wie erwähnt, scheinen bereits in der ersten Sequenz der sprachlichen Äußerung zentrale prosodische Markierungen enthalten zu sein, die zeitig eine Differenzierung der einzelnen Sprechfassungen voneinander erlauben, was man auch im Spektrogramm in Abb. 14 a-c gut dargestellt findet.

**Sprechfassung I**

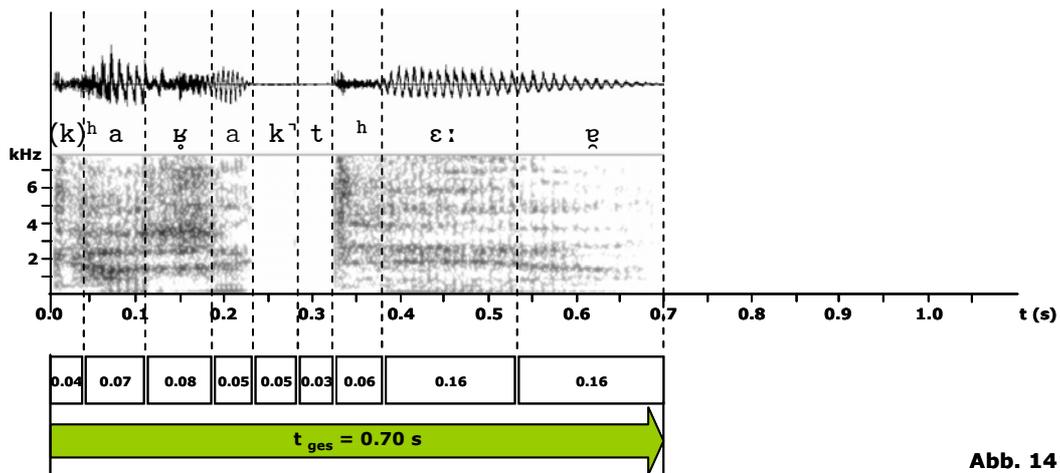


Abb. 14 a

**Sprechfassung II**

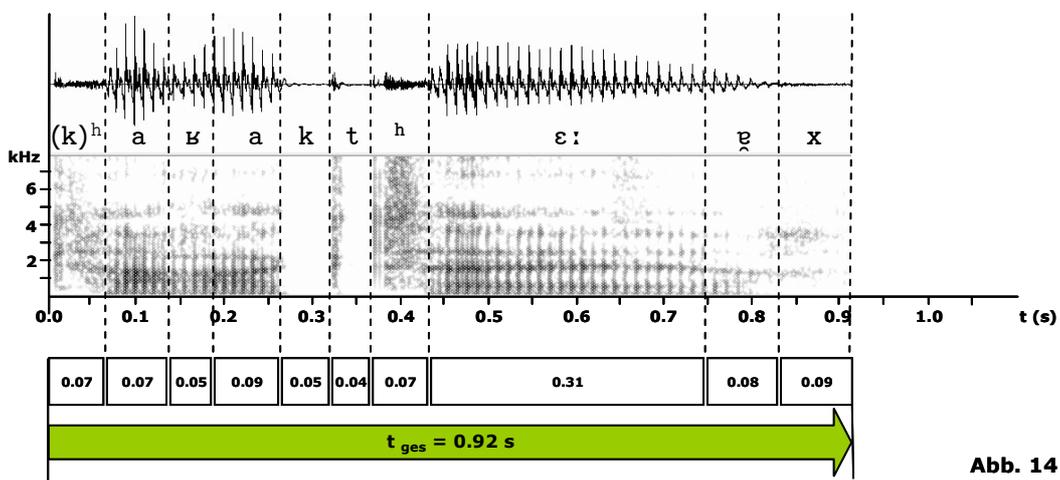


Abb. 14 b

**Sprechfassung III**

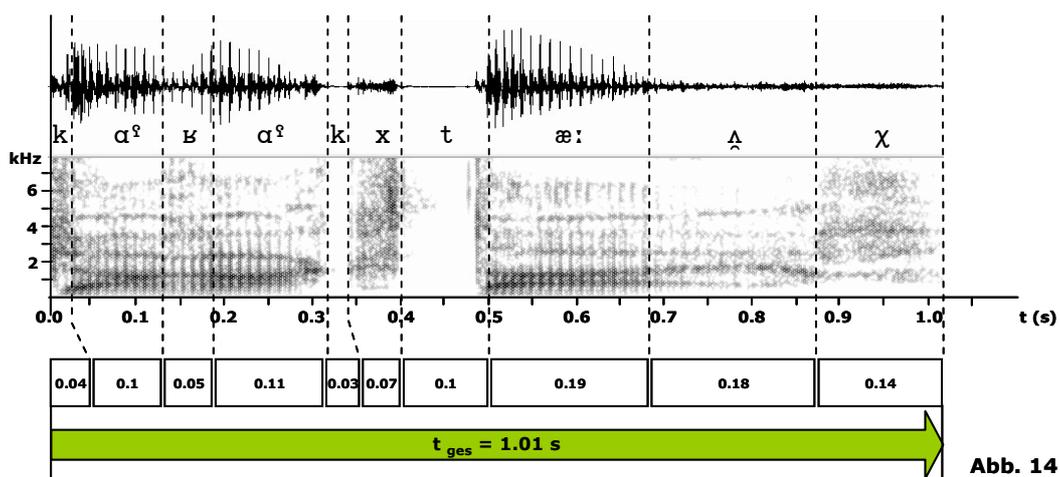


Abb. 14 c

Abb. 14 a-c: Temporale Struktur und Lautsegmentierung von „Caractère“ aller drei Sprechfassungen im Vergleich. Alle Angaben in Sekunden (s).

Diese Markierungen liegen zum großen Teil in segmentalen Merkmalen, die auch bestimmte temporale Erscheinungen mit sich bringen. Auffällig scheinen zunächst die verlängerten Silbencodas in SF II und III (Abb. 14 b-c) der jeweils dritten Silbe, um den velaren [x] bzw. postvelaren Frikativ [χ], der in SF III sowohl durch seine temporale Ausdehnung, als auch durch seine starke Friktion dominiert. In Tabelle 12 ist die temporale Struktur der einzelnen Laute in den verschiedenen SF dargestellt. Die Akzentsilbe ist dabei jeweils grau unterlegt.

### Caractère.

<b>SF I</b>	k <sup>h</sup>	a	ɣ	a	k <sup>ʳ</sup>	t	h	ɛ:	ɸ	
	0.04	0.07	0.08	0.05	0.05	0.03	0.06	0.16	0.16	
<b>SF II</b>	k <sup>h</sup>	a	ɣ	a	k	t	h	ɛ:	ɸ	x
	0.07	0.07	0.05	0.09	0.05	0.04	0.07	0.31	0.08	0.09
<b>SF III</b>	k	ɑː	ɣ	ɑː	k	x	t	æ:	ʌ	χ
	0.04	0.1	0.05	0.11	0.03	0.07	0.1	0.19	0.18	0.14

**Tab. 12: Temporale Struktur von „Caractère“ aller drei Sprechfassungen im Überblick. Alle Angaben in Sekunden. Die Markierung hebt die Akzentsilbe hervor.**

Aus der Tabelle ergibt sich zunächst die Abgrenzung des Vorakzentintervalls (die ersten zwei Silben) von der Akzentsilbe. Über alle SF hinweg erstreckt sich das Vorakzentintervall, obwohl es aus zwei Silben besteht, über einen kürzeren Zeitraum als die Akzentsilbe, wobei die Unterschiede in SF I mit 0.29s vs. 0.41s nicht so stark ausgeprägt sind.

Die Silben innerhalb der SF unterscheiden sich maßgeblich in ihrer Vokalqualität und auch Vokalquantität. Die Vokalquantität wiederum ist eng an temporale Strukturen gebunden und stellt ein wichtiges differenzierendes Merkmal der einzelnen SF dar. Während das Vorakzentintervall über die SF hinweg relativ stabil bleibt (SF I = 0.29s, SF II = 0.33s und SF III = 0.40s), manifestieren sich in der Akzentsilbe deutliche Variationen zwischen den SF, die von SF I mit 0.41s, über SF II mit 0.59s und SF III mit 0.61s stetig ansteigen. Diese Angaben decken sich aber nicht genau mit den Proportionen der Akzentsilbe zu der gesamten Sequenz (Caractère). Dabei nimmt die Akzentsilbe folgende Anteile ein: in SF I = 58.6%, in SF II = 63.0% und in SF III = 60.4%. Die Werte

verdeutlichen, dass die temporale Ausdehnung der Akzentsilbe im Vergleich zum Vorakzentintervall in SF II besonders ausgeprägt ist.

Wenn man innerhalb der Akzentsilbe die Realisierung des Diphthongs genauer betrachtet, so lässt sich den Werten aus Tabelle 12 entnehmen, dass sowohl in SF I als auch in SF III beide Vokalbestandteile in ihrer temporalen Beschaffenheit nahezu gleich lang realisiert werden (in SF I 0.16s vs. 0.16s und SF III 0.19s vs. 0.18s), wohingegen in SF II der zweite Vokalteil [ɥ] mit 0.08s gerade ein Viertel des ersten Vokalbestandteils [ɛ:] mit 0.31s einnimmt<sup>11</sup>.

Aus den temporalen Strukturen (Tab. 12) lassen sich auch Aussagen über die Markierung von Nebenakzenten ableiten. Verglichen mit dem Akzentvokal, ordnen sich in SF I und II die Nebenakzente, die auf den Vokalen /a/ von „C a r a c t è r e“ gesetzt sind, nahezu unter und sind als solche in ihrer temporalen Ausdehnung kaum zu perzipieren. Im Vergleich dazu zeichnet sich SF III durch dominante Nebenakzente auf den pharyngalisierten Vokalen [ɑ̠] aus, die mit 0.10s und 0.13s aus dem lautlichen Umfeld stärker heraustreten als in den anderen beiden SF. Dieses Phänomen erklärt auch, warum die Akzentsilbe im Vergleich zur Gesamtzeit der Sprechsequenz mit 60.4% anteilig weniger Zeit umfasst als die vergleichbare Stelle in SF II mit 63.0% (s.o.).

Letztlich unterscheiden sich die Sprechweisen des Parfumnamens auch unter segmentalen Gesichtspunkten, die sich im Spektrogramm nachweisen lassen (vgl. Abb. 14 a-c). Dabei handelt es sich allerdings um keine phonologisch relevanten Kategorien, sondern um die Laute modifizierende artikulatorische Nebenerscheinungen, die von daher auch der Prosodie zugesprochen werden können. Am deutlichsten setzt sich SF III von SF I und SF II ab, was sich zunächst in den stark pharyngalisierten /a/-Realisationen [ɑ̠] manifestiert, die sich regressiv auf die Bildungsstelle des vorausgehenden Plosivs übertragen. Des Weiteren zeichnet sich SF III neben der fehlenden Aspiration der Explosive zusätzlich durch eine starke Friktion des postvelaren Frikativs [χ] aus. Die damit verbundene Stimmlosigkeit überträgt sich regressiv auf den vorausgehenden Vokalteil [ɑ̠] und entstimmt auch ihn nahezu vollständig.

Die anderen beiden SF unterscheiden sich in erster Linie durch die Stimmqualität. Während SF I über die gesamte Sequenz starke Behauchung der Stimme

---

<sup>11</sup> Als Kriterien des Segmentierens der einzelnen Diphthongbestandteile wurden auditiv-visuelle Veränderungen als Richtlinie angelegt.

aufweist (vgl. hohe Frequenzbereiche in Abb. 14 a), zeichnet sich SF II durch einen sehr hohen stimmhaften Anteil aus (vgl. Formantstruktur in Abb. 14 b). Im folgenden Abschnitt soll die Zwischensequenz „M a c h t a m M a n n“ unter temporalen Gesichtspunkten genauer analysiert werden. Die temporale Struktur dieser Sequenz steht, wie in Tabelle 13 zusammengefasst, im Verhältnis zur Dauer der gesamten Äußerung, d.h. SF I (0.66s) ist kürzer als SF II (0.84s) und SF III (1.00s). In ihrer Binnenstruktur zeigen sich aber Differenzen, die sich auf Dauerunterschiede einzelner Laute und bestimmter Lautverbindungen beziehen.

### **M a c h t a m M a n n .**

<b>SF I</b>	m	a	χ	t	ǣ	m:	a	n
	0.04	0.05	0.06	0.06	0.06	0.13	0.11	0.15
<b>SF II</b>	m	a	χ	t	ǣ	m:	a	n
	0.05	0.08	0.07	0.09	0.08	0.14	0.15	0.18
<b>SF III</b>	m	a	ʀ	t	a	m:	a	n
	0.08	0.08	0.09	0.1	0.1	0.14	0.15	0.26

**Tab. 13: Temporale Struktur von „M a c h t a m M a n n“ aller drei Sprechfassungen im Überblick. Alle Angaben in Sekunden.**

Abbildung 15 a-c zeigt das Spektrum mit den Lautsegmentierungen innerhalb der einzelnen Realisationen grafisch. Aus der Darstellung lassen sich neben Aussagen zur Lautquantität auch Aussagen bezüglich der Qualität einzelner Laute ableiten. Die Äußerungssequenz besteht aus den zwei von der Anzahl der Laute her identischen Teilen [m a χ t] und [ǣ m : a n], wobei in allen drei SF die zweite Gruppe ([ǣ m : a n]) mit 68% für SF I und jeweils 65% für die SF II und SF III, eine dominierende Stellung einnimmt und somit mehr Spielraum für temporale Variabilität bietet. Das Bindeglied der beiden Lautgruppen [m a χ t ǣ m : a n] stellt die Folge des alveolaren Plosives, der jeweils ohne Aspiration und des folgenden Vokals dar, der anstelle eines festen Stimmeinsatzes mit Knarrstimme realisiert wird (vgl. Abb. 15 a-c).

## Sprechfassung I

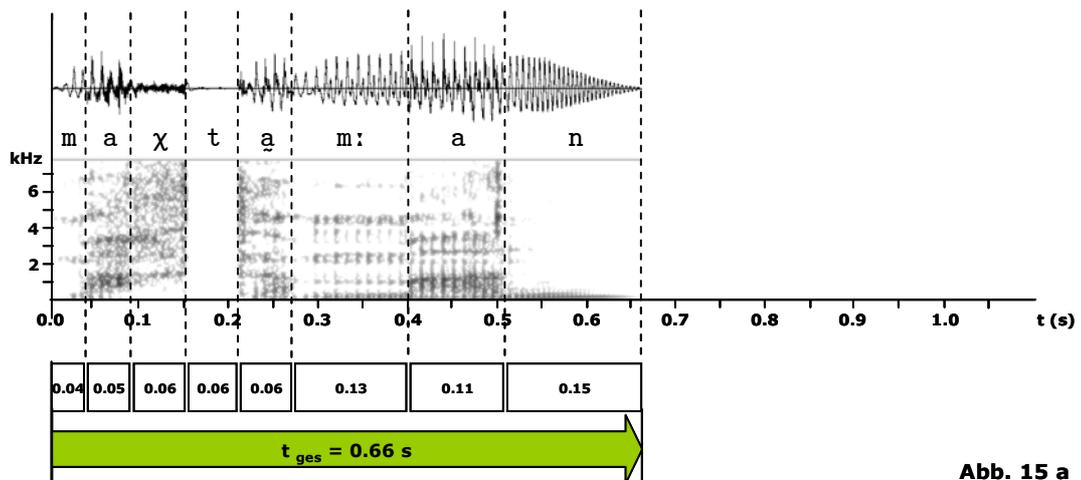


Abb. 15 a

## Sprechfassung II

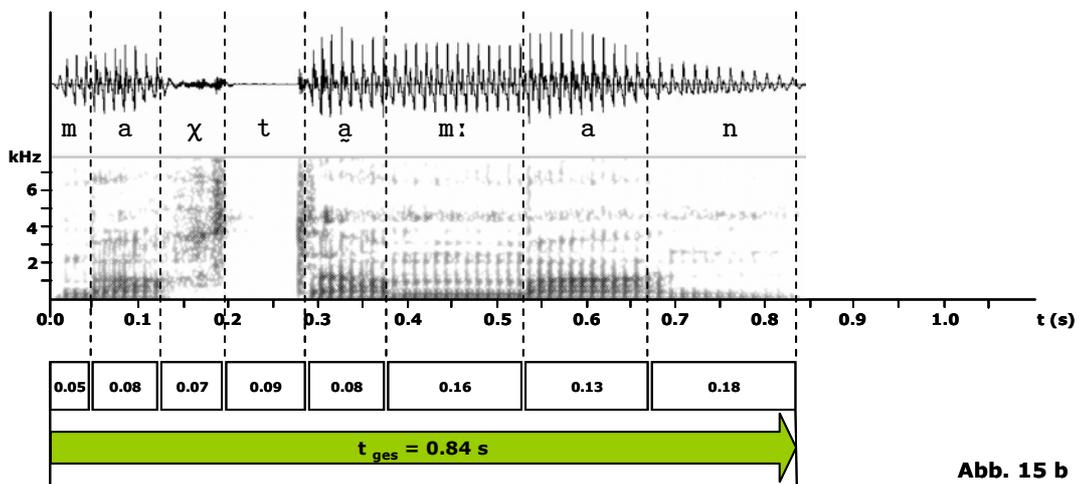


Abb. 15 b

## Sprechfassung III

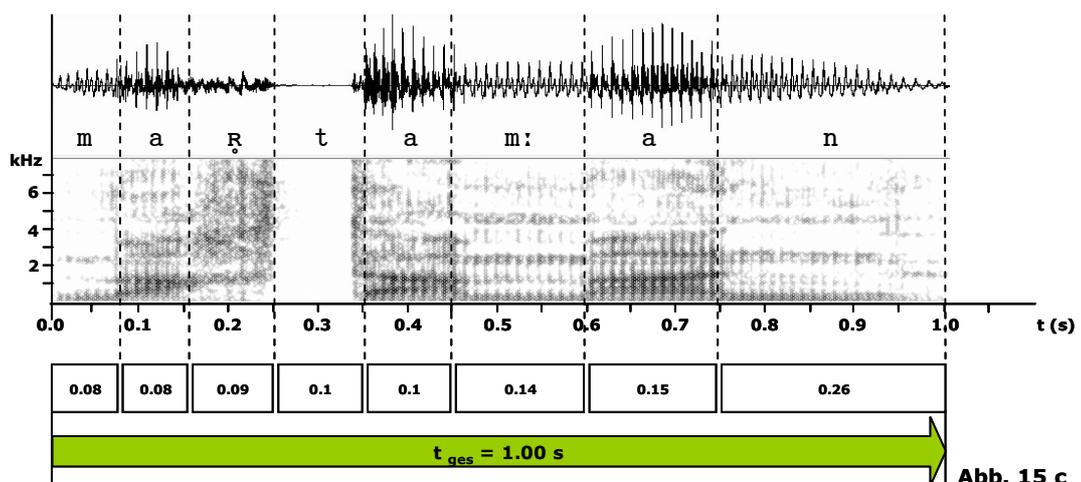


Abb. 15 c

Abb. 15 a-c: Temporale Struktur und Lautsegmentierung von „Macht am Mann“ aller drei Sprechfassungen im Vergleich. Alle Angaben in Sekunden (s).

Auf diese Weise sind die beiden Lautgruppen eng miteinander, ohne jeden Neueinsatz der Stimme, verschmolzen. Den Werten von Tab. 13 lässt sich darüber hinaus entnehmen, dass der alveolare Plosiv [t] und das darauf folgende [ä] sich jeweils durch nahezu identische Lautdauern, die sich aber über die SF hinweg verändern, auszeichnen (SF I 0.06s vs. 0.06s, SF II 0.09s vs. 0.08s, SF III 0.1s vs. 0.1s).

Die beiden Lautgruppen unterscheiden sich auffällig in der Stimmqualität. Während das „macht“ über die SF hinweg in Modalstimme realisiert wird, lassen sich in „am Mann“ Knarranteile unterschiedlicher Ausprägung nachweisen (vgl. Abb. 15 a-c), die auf diese Weise zugleich die Abgrenzung der beiden Gruppen voneinander durch die fehlende Pause kompensieren.

In der Betrachtung einzelner Laute lassen sich die größten Dauerunterschiede im [n] von „am Mann“ nachweisen. Hier variiert die Realisierung innerhalb der SF um 0.11s von SF I mit 0.15s, über SF II mit 0.18s zu SF III mit 0.26s. In allen drei Varianten nimmt das [n] damit eine größere Lautdauer als der vorausgehende Vokal ein. Während in SF I und SF II die Unterschiede vernachlässigbar gering sind, stellt die fast verdoppelte Lautdauer des Nasalkonsonanten im Vergleich zum Vokal (0.15s vs. 0.26s) in SF III eine Akzentverschiebung, die ihre Ursache in der temporalen Beschaffenheit der Laute hat, dar.

Es bleibt abschließend unter Gesichtspunkten der Temporalität und der Artikulation noch die Betrachtung der Schlusssequenz „das Markante am Mann“, die in Abbildung 16 a-c grafisch dargestellt ist.

Auch die dritte Äußerungssequenz verhält sich, ähnlich wie die zweite, in ihrer temporalen Ausdehnung im Verhältnis zur gesamten Äußerung. So erweist sich die sprecherische Realisierung von SF II mit 1.25s länger als die von SF I mit 1.13s, aber kürzer als jene von SF III mit 1.57s. Zwischen SF I und SF III liegt mit 0.44s eine Differenz von fast einer halben Sekunde, was für die Kürze der Sequenz deutliche temporale Unterschiede mit sich bringt. Durch die Angabe der temporalen Ausdehnung in Silben pro Sekunden (S/s) lässt sich dieser Fakt besonders deutlich darstellen. Während in SF I durchschnittlich 5.31S/s realisiert werden, vermindern sich die Werte in SF II mit 4.8S/s und SF III mit 3.82S/s nachweislich.

**Sprechfassung I**

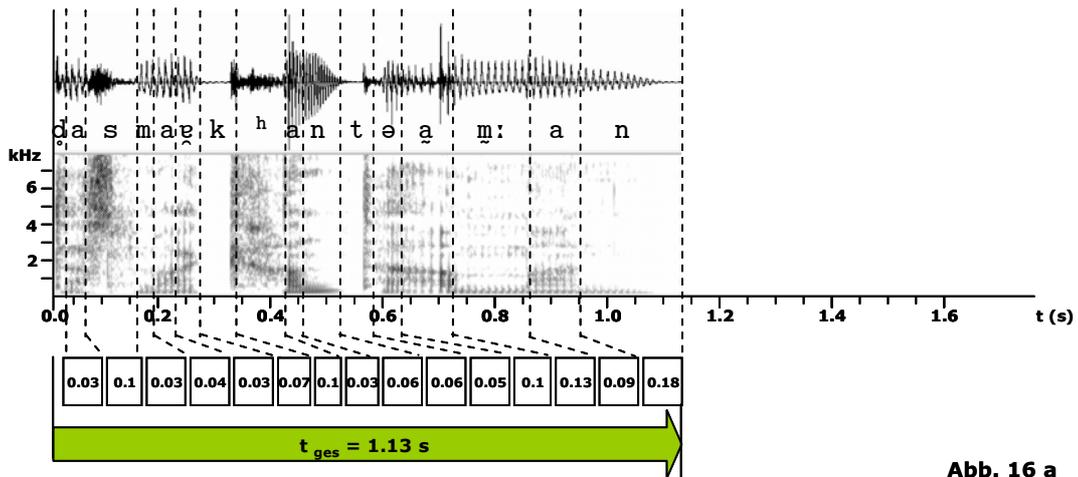


Abb. 16 a

**Sprechfassung II**

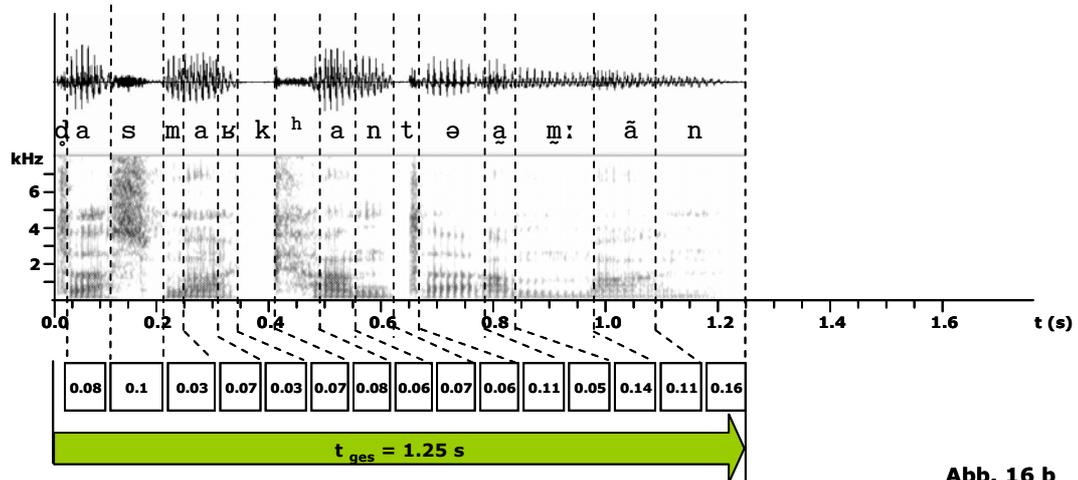


Abb. 16 b

**Sprechfassung III**

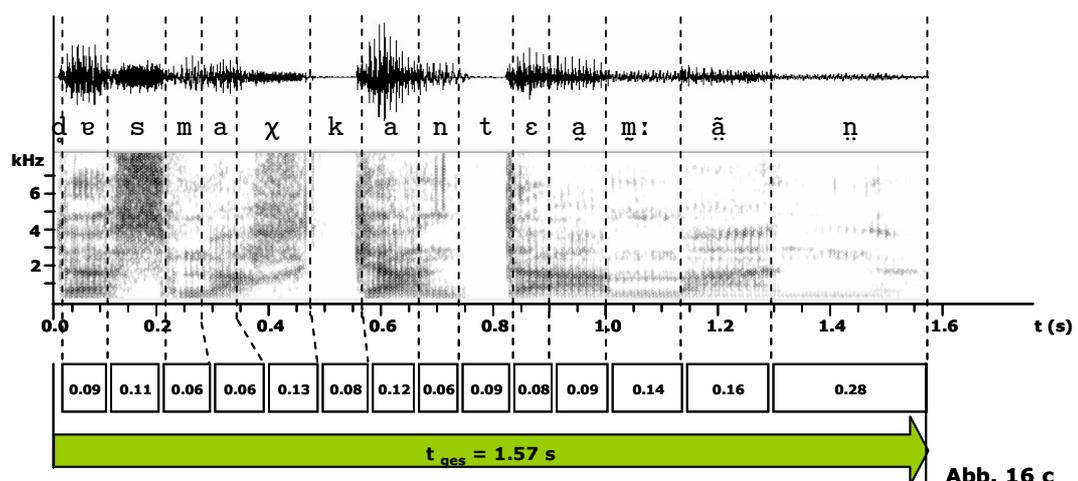


Abb. 16 c

Abb. 16 a-c: Temporale Struktur und Lautsegmentierung von „das Markante am Mann“ aller drei Sprechfassungen im Vergleich. Alle Angaben in Sekunden (s).

Schwieriger als in den vorausgegangenen Sequenzen gestaltet sich allerdings in diesem Sprechabschnitt die Segmentierung einzelner Laute an bestimmten Übergängen (z.B. **das Markante am Mann**), weil nicht immer zweifelsfrei entschieden werden kann, welche koartikulatorische Erscheinung (Velumstätigkeit) welchem Laut zugeschlagen wird. So stellt sich beispielsweise bei dem Übergang vom alveolaren stimmlosen Frikativ [s] in den stimmhaften Nasal [m] die Frage, ob sich auch Stimmlosigkeit auf den folgenden Nasal in Form von [s̥m̥] überträgt, oder ob Nasalität regressiv auf den Frikativ ([s̥m]) übertragen wird, oder aber, ob es sich um einen reziproken Prozess handelt, bei dem [s̥m̥] als Resultat angenommen wird. Um die lautliche Segmentierung in jedem Fall konsistent zu halten (vgl. Simpson 1998:29ff), wurde großer Wert auf Beständigkeit gelegt, indem beispielsweise für den Übergang vom Frikativ in den Nasalkonsonanten die Stimmhaftigkeit als Kriterium angelegt und vollständig dem Nasalkonsonanten zugesprochen wurde. Damit bildet sie zugleich die Grenze zwischen den benachbarten Lauten (vgl. Abb. 16 a-c). Als Leitlinien für die Konsistenz in der Segmentierung wurden für alle SF einheitlich auditive und auch visuelle Kriterien zugrunde gelegt (vgl. Kap. 6.2).

Weniger aufschlussreich als in den vorangegangenen Sequenzen gestaltet sich die Binnenstruktur von „das Markante am Mann“, wie die Werte in Tab. 14 verdeutlichen.

Anhand der Markierung lässt sich aber entnehmen, dass die sprecherische Umsetzung des /r/ in „Markante“ in jeder SF unterschiedlich realisiert wird. Während in SF I ein vokalisiertes [ɐ̯] gesprochen wird, enthalten SF II und SF III einen postvelaren Frikativ als /r/-Variante. Die Realisationen des /r/ in SF II und III unterscheiden sich demnach nicht in ihrem artikulierenden Organ und ihrer Artikulationsstelle, dafür in Bezug auf ihre Stimmbeteiligung (vgl. Abb. 16 b-c). Während das [ɐ̯] in SF II durchgängig mit Stimmbeteiligung realisiert wird, hebt sich das [χ] in SF III, was ohne jede Stimmbeteiligung und mit starker Friktion realisiert wird, deutlich davon ab. Zudem umfasst das [χ] in SF III mit 0.13s mehr als das Vierfache der /r/-Realisationen aus SF I und II, deren Ausdehnung lediglich bei 0.03s liegt, und dominiert damit auch sein lautliches Umfeld.

**das Markante am Mann.**

	d̥	a	s	m	a	g	k	<sup>h</sup>	a	n
<b>SF I</b>		0.03	0.1	0.03	0.04	0.03	0.07	0.1	0.03	0.06
	t	e		ɑ	ɯ:	a	n			
	0.06	0.05		0.1	0.13	0.09	0.18			
<b>SF II</b>	d̥	a	s	m	a	ɣ	k	<sup>h</sup>	a	n
		0.08	0.1	0.03	0.07	0.03	0.07	0.08	0.06	0.07
	t	e		ɑ	ɯ:	ã	n			
	0.06	0.11		0.05	0.14	0.11	0.16			
<b>SF III</b>	d̥	ɛ	s	m	a	χ	k	a	n	
		0.09	0.11	0.06	0.06	0.13	0.08	0.12	0.06	
	t	ɛ		ɑ	ɯ:	ã	ɲ			
	0.09	0.08		0.09	0.14	0.16	0.28			

**Tab. 14: Temporale Struktur von „das Markante am Mann“ aller drei Sprechfassungen im Überblick. Alle Angaben in Sekunden. Die Markierung hebt die unterschiedlichen r-Realisationen hervor.**

Aus Abbildung 16 a-c lassen sich auch Aussagen in Bezug auf die Aspiration des velaren Plosivs [k] ableiten. Dieser ist in SF I und SF II deutlich aspiriert (vgl. Abb. 16 a-b). In beiden Fällen überträgt sich die Aspiration koartikulatorisch auf den folgenden Vokal [a], der darauf hin in SF I zu nahezu zwei Dritteln entstimmt ist. In SF II lässt sich allenfalls ein verspäteter Stimmeinsatz des Vokals aufzeigen. Im Gegensatz dazu zeichnet sich SF III, wie bereits für die ersten beiden Sprechsequenzen beschrieben, wieder durch die völlige Abstinenz der Aspiration des velaren Plosivs aus. Diese intendierte Interferenzerscheinung hat zur Folge, dass SF III, in der ein französischer Akzent imitiert wird, auch als französisch klingend erkannt und empfunden wird.

Wie bereits für die vorausgegangenen Passagen erwähnt, zeichnet sich SF I im Vergleich zu SF II und SF III durch einen hohen Anteil an behauchter Stimme aus. Das lässt sich beispielhaft an der Schnittstelle des velaren Plosivs [k] zum Vokal [a] in „das Markante am Mann“ belegen. Vergleicht man die Realisationen in SF I und SF II miteinander, in denen jeweils der stimmlose Plosiv von deutlicher Aspiration begleitet ist, so ergeben sich rechnerisch aus der Formel  $\bar{a} = a : a + ^h$  Werte für die Anteile der Stimmhaftigkeit, die für SF I mit 23% deutlich niedriger liegen als für SF II mit 43%, und die damit bestätigen,

dass sich SF II durch eine größere Stimmhaftigkeit auszeichnet, wohingegen SF I stärkere Hauchanteile in der Stimme aufweist.

Abschließend sollen noch einige Bemerkungen zur Stimmqualität gemacht werden, die sich wie bereits in Sequenz zwei besprochen auf den Einsatz von Knarrstimme beziehen. Über alle SF hinweg wird Knarren an Stelle des festen Vokaleinsatzes in „**a** m M a n n“ eingesetzt und geht nahtlos in den folgenden Nasalkonsonanten über. In SF III ändert sich noch zusätzlich die Stimmqualität innerhalb der letzten beiden Segmente [mǎ̃ñ] was vermutlich auf die extrem tiefe Stimmlage am Ende der Äußerung zurückzuführen ist, und was dazu führt, dass die Lösungstiefe nahezu völlig entstimmt und dadurch von einer behauchten Stimmqualität begleitet wird. Dieses Phänomen unterscheidet SF III von den anderen beiden SF, die aus der Knarrstimme wieder zurückgehen und die Äußerung als Ganzes in Modalstimme beenden.

In Bezug auf die Stimmqualität soll abschließend erwähnt sein, dass sich sowohl in SF II als auch in SF III, koartikulatorisch bedingt durch die beidseitige Nachbarschaft von Nasalkonsonanten, in [mǎ̃ñ] eine Nasalisierungstendenz des Vokals in internasaler Position ergibt.

Zusammenfassend kann festgehalten werden, dass die temporale Komponente, begleitet von artikulatorischen Merkmalen, bezüglich der Diskriminierung der drei SF voneinander sowohl in ihrer äußeren Struktur ( $t_{ges}$ ), als auch in ihrer Binnenstruktur (Pausen, Lautdauern und Lautverbindungen) einen entscheidenden Einfluss ausübt. Vor allem SF III hebt sich unter den besprochenen Gesichtspunkten der Temporalität und der Artikulation deutlich von SF I und von SF II ab, die sich vornehmlich durch den Anteil der Stimmbeteiligung unterscheiden.

Im folgenden Abschnitt sollen nun melodische Veränderlichkeiten in den einzelnen SF genauer beschrieben werden. Dabei geht es ebenfalls darum, Tendenzen aufzuzeigen, die es ermöglichen die drei SF mit ihren individuellen Verlaufskriterien so zu beschreiben, dass sie sich hinreichend voneinander unterscheiden lassen.

## **Melodische Veränderlichkeit**

Um die Konturen der Melodiekurve aussagekräftig beschreiben zu können, wurden 14 Messpunkte innerhalb der Äußerung definiert, an denen zum einen

die  $F_0$ -Werte in Hz und zum anderen die entsprechenden Tonhöhen in Halbtonschritten festgehalten wurden. Diese Messpunkte entsprechen repräsentativ ausgewählten Punkten für die Beschreibung der Melodiekurve und wurden zumeist in den Vokalen bzw. den sich anschließenden Nasalkonsonanten festgelegt, so dass sich folgendes Schema ergibt „Caractère | Macht am Mann | das Markante am Mann.“

Da wie in Kapitel 4.1 beschrieben, die ermittelten  $F_0$ -Werte nicht linear mit dem menschlichen Höreindruck korrelieren, werden die Melodiekurven zudem in Halbtonschritten, wie sie in der Notenlehre verwendet werden, dargestellt, weil dieses Notationssystem einer für unsere Hörgewohnheiten linearen Ordnung entspricht. Die verschiedenen Abbildungen werden zeigen, dass sich hinter scheinbar kleinen Bewegungen der  $F_0$ -Werte oft deutlich hörbare Intervallsprünge verbergen, was auf die tiefe, im Bassbereich angesiedelte, Sprechstimmlage des Sprechers zurückzuführen ist.

Auch unter Gesichtspunkten der Melodisierung differieren die einzelnen SF deutlich voneinander. Während sich aber in Bezug auf die Temporalität und die Artikulation, wie oben beschrieben, SF III merklich von SF I und von SF II absetzt, liegen diese in ihrem melodischen Verlauf und in ihrer melodischen Variabilität nicht so weit voneinander entfernt. Es wird sich zeigen, dass aber mit Augenmerk auf die melodische Komponente SF I spezifische Merkmale aufweist und sich somit von SF II und SF III hörbar abhebt.

In den Abbildungen 17 a-c sind die Melodiekurven für die Realisierung von „Caractère“ dargestellt. Dabei wurden die Messwerte jeweils in der Mitte der einzelnen Vokale bzw. Nasalkonsonanten ermittelt.

Im Befund (Abb. 17 a-c) zeigen sich drei Konturen, die zunächst über die SF hinweg gemeinsam haben, dass die ersten beiden Messpunkte gegenüber den letzten beiden erhöht liegen, so dass ein terminales Melodiemuster realisiert wird, das damit die Sprechsequenz „Caractère“ bereits als in sich abgeschlossen erscheinen lässt.

Zudem kann in den drei Abbildungen (17 a-c) auch noch einmal die Verschiedenartigkeit der Realisierung der Silbenkudas aufgezeigt werden, die in SF II und SF III jeweils um den velaren bzw. postvelaren Frikativ verlängert sind. In SF III wird durch die starke temporale Dehnung des Frikativs der stimmhafte Anteil des vorausgehenden Diphthongs im Vergleich zu den anderen beiden Sprechfassungen stark verkürzt.

**Sprechfassung I**

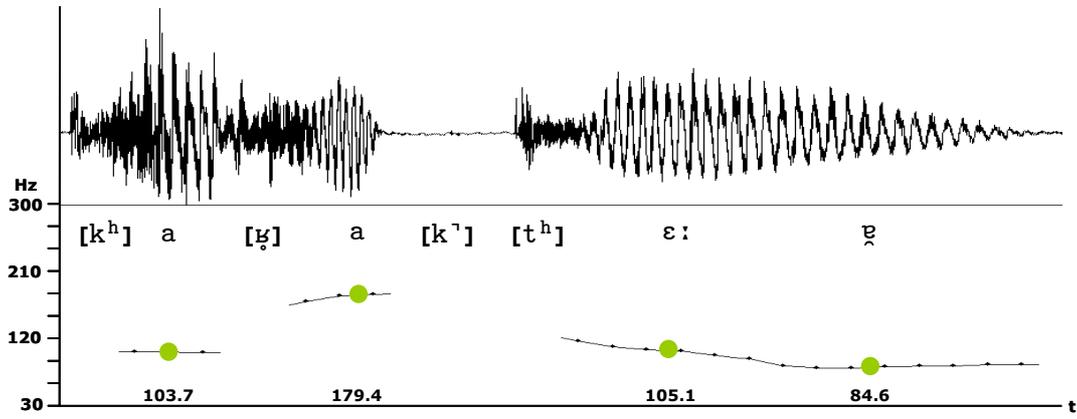


Abb. 17 a

**Sprechfassung II**

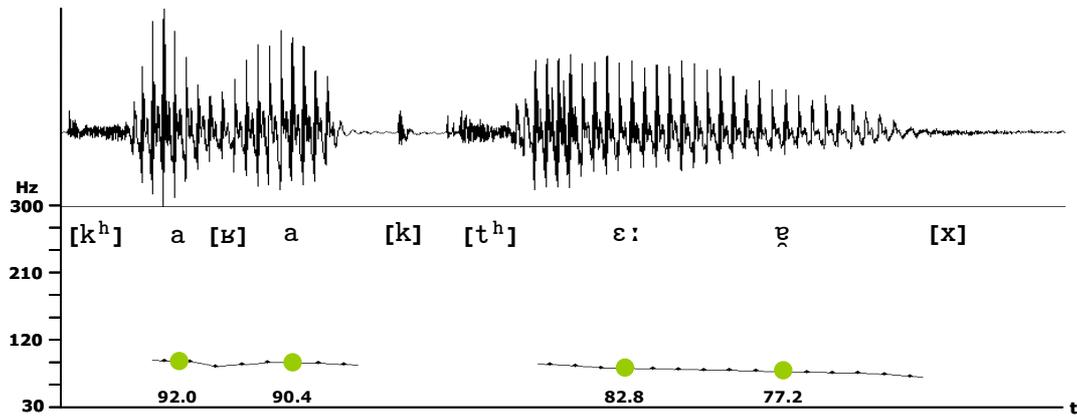


Abb. 17 b

**Sprechfassung III**

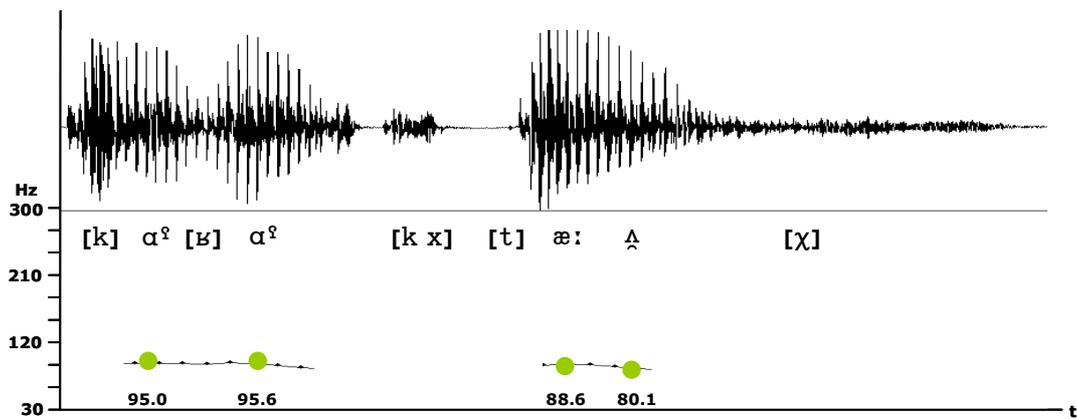
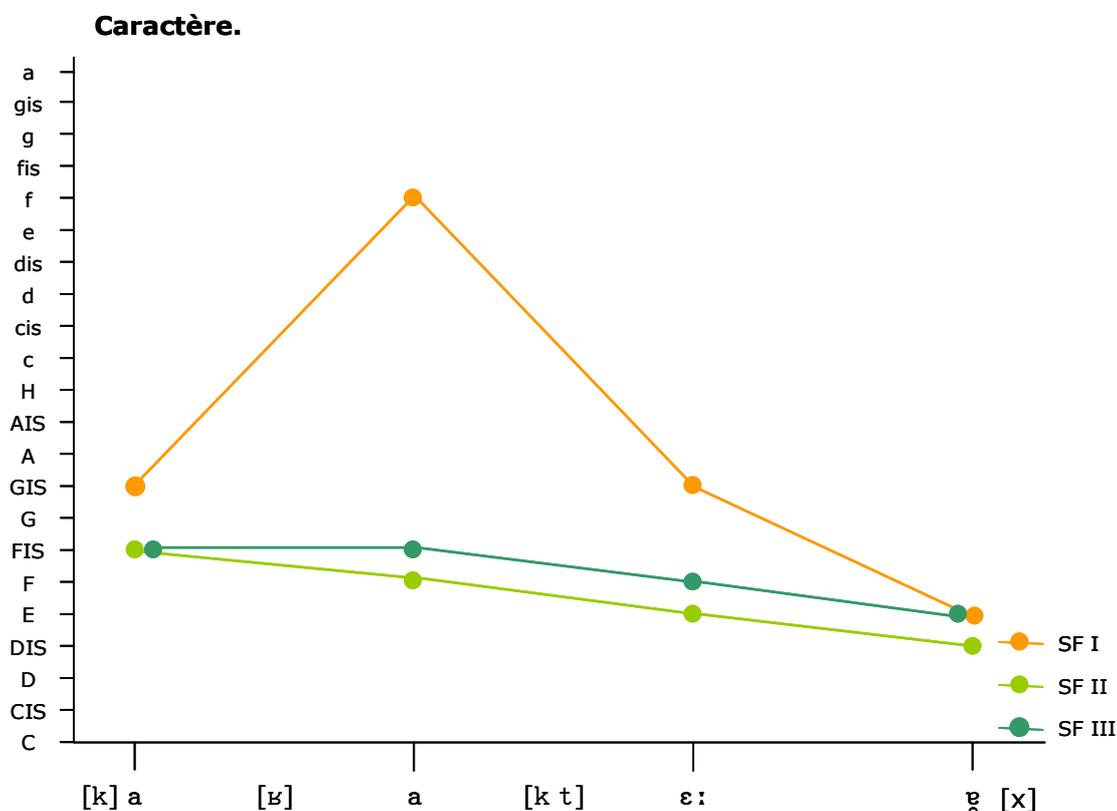


Abb. 17 c

**Abb. 17 a-c: Grundfrequenzverlauf von „Caractère“ aller drei Sprechfassungen im Vergleich. Alle Angaben in Hz.**

Während sich in SF II und III die Grundfrequenzkurve nur minimal verändert, kann man in SF I deutliche Bewegungen sehen, die in Abbildung 18 noch einmal in Halbtonschritten dargestellt sind.



**Abb. 18: Darstellung des melodischen Verlaufs von „Caractère“ in Halbtonschritten in allen drei Sprechfassungen im Vergleich.**

Die Kurven verdeutlichen, dass in SF I und SF III ein melodischer Gipfel auf der zweiten Silbe entsteht, der sich in SF III jedoch nur durch einen Halbton aus seinem Umfeld heraushebt, wohingegen sich in SF I zwischen dem melodischen Akzent und dem Umfeld ein Intervall von 9 Halbtönen (große Sexte) ergibt. Dieser Intervallsprung von GIS zu f in SF I zieht einen Registerwechsel von der Bruststimme in die Kopfstimme nach sich, und lässt damit den Akzent aus dem Vor- und Nachakzentintervall deutlich heraustreten<sup>12</sup>. Zudem liegt SF I im gesamten Verlauf im Vergleich zu den anderen SF leicht erhöht. Mit einem Intervall von 13 Halbtonschritten (kleine None) zwischen den Extrempunkten

<sup>12</sup> Die Einordnung der Register erfolgte nach auditiven Kriterien. Terminologisch werden unterschiedliche Begriffe verwendet: Sundberg (1997:74f) spricht für die Singstimme von Modal- und Falsettregister; Wirth (1995:128f) verwendet die Begriffe Brust- und Kopfstimme (vgl. auch Kap. 3.3.2).

(Messpunkt 2 und Messpunkt 4) setzt sich SF I in seiner melodischen Variabilität deutlich von SF II und SF III ab, die mit einem Halbton Differenz von Messpunkt 2 bis Messpunkt 4 parallel verlaufen, wobei SF II kontinuierlich unter SF III liegt. Beide SF werden durchgängig im Brustregister realisiert. Unterschiede zwischen SF II und SF III finden sich lediglich in der melodischen Vorverlagerung des Akzents in SF II auf die erste Silbe. Von da an fällt die Melodiekurve in Halbtonschritten ab, so dass ein Intervall von 3 Halbtonschritten (kleine Terz) zwischen den Extrempunkten (Messpunkt 1 und Messpunkt 4) entsteht. In Tabelle 15 sind die unterschiedlichen Werte noch einmal zusammengefasst.

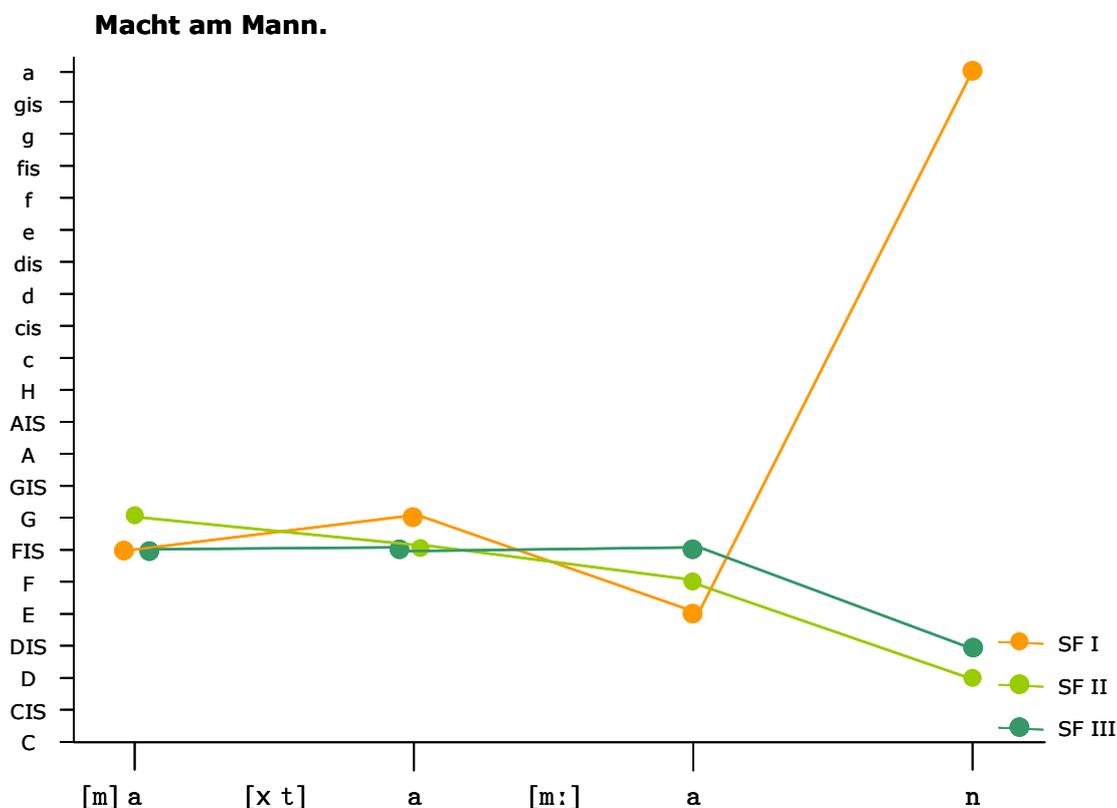
### Caractère.

	[k <sup>h</sup> ]	a	[ɣ]	a	[k <sup>ʀ</sup> ]	[t <sup>h</sup> ]	ɛ:	ɐ̃		U <sub>I</sub>
<b>SF I</b>		103.7		179.4			105.1	84.6		
		GIS		f			GIS	E		
<b>Intervall</b>			9		9			4		13
			HTS		HTS			HTS		HTS
			↑		↓			↓		
	[k <sup>h</sup> ]	a	[ɣ]	a	[k]	[t <sup>h</sup> ]	ɛ:	ɐ̃	[x]	U <sub>I</sub>
<b>SF II</b>		92.0		90.4			82.8	77.2		
		FIS		F			E	DIS		
<b>Intervall</b>			1		1			1		3
			HTS		HTS			HTS		HTS
			↓		↓			↓		
	[k]	ɑ̃	[ɣ]	ɑ̃	[kx]	[t]	æ:	ʌ	[χ]	U <sub>I</sub>
<b>SF III</b>		95.0		95.6			88.6	80.1		
		FIS		FIS			F	E		
<b>Intervall</b>			0		1			1		2
			HTS		HTS			HTS		HTS
					↓			↓		

**Tab. 15: Melodieverlauf von „Caractère“ aller drei Sprechfassungen im Überblick. Alle Angaben in Hz und Halbtonschritten. Die grau unterlegte Spalte gibt das Intervall zwischen den Extrempunkten an.**

Die Zeilen, in denen die Intervalle abgetragen sind und die grau unterlegte Spalte, die das Gesamtintervall zwischen den Extrempunkten angibt, verdeutlichen, dass die melodische Variabilität in SF I sich sowohl in der Qualität als auch der Quantität der Melodiesprünge deutlich von SF II und SF III absetzt. Diese Tendenz lässt sich auch in der sich anschließenden Äußerungssequenz „Macht am Mann“ beobachten. Auch hier liegen SF II und SF III in ihrem melodischen Verlauf dichter beieinander und unterscheiden sich hörbar von SF

I. In Abbildung 19 ist die Kurve in Halbtonschritten für diese Zwischensequenz dargestellt.



**Abb. 19: Darstellung des melodischen Verlaufs von „Macht am Mann“ in Halbtonschritten in allen drei Sprechfassungen im Vergleich.**

Aus Abbildung 19 lässt sich ablesen, wie annähernd parallel die Kurvenverläufe aller drei SF im Vorakzentintervall bis Messpunkt 3 gestaltet sind. SF I zeigt bis dahin nur eine geringfügig bewegtere Kurve. Mit der Tieferlegung des Akzentes im Vergleich zum Vor- und Nachakzentintervall unterscheidet sich SF I aber von den anderen SF. Zudem fällt das extrem große Intervall von 17 Halbtonschritten auf, was auf engstem Raum in der Gleitbewegung vom Vokal auf den sich anschließenden Nasalkonsonanten zwischen Messpunkt 3 und 4 realisiert wird. Diese Verläufe sind in Abbildung 20 a-c noch einmal als  $F_0$ -Kurven dargestellt.

Dem Anstieg der Grundfrequenz von 84.1Hz auf 218.0Hz zwischen Messpunkt 3 und 4 in Abbildung 20 a liegt ein Intervall von über einer großen Dezime von E nach a zu Grunde, was neben einem deutlichen Registerwechsel von der Bruststimme in die Kopfstimme gleichzeitig zu einer kurzzeitigen Verlegung

**Sprechfassung I**

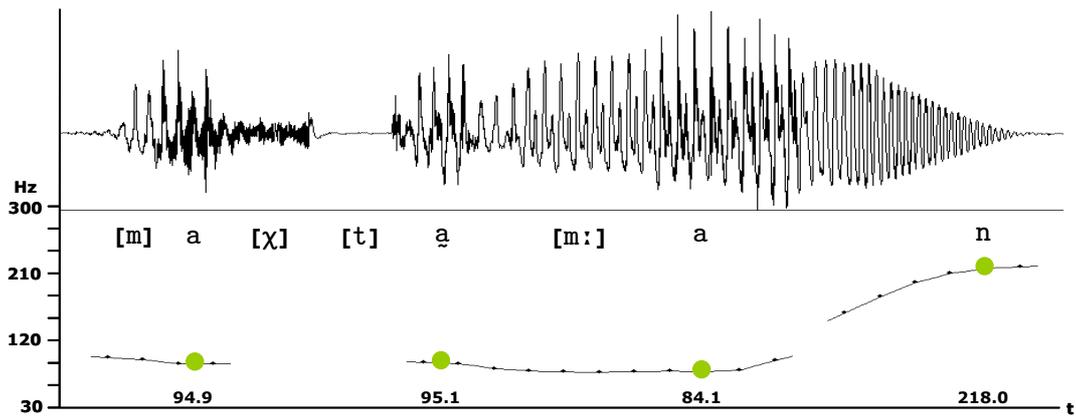


Abb. 20 a

**Sprechfassung II**

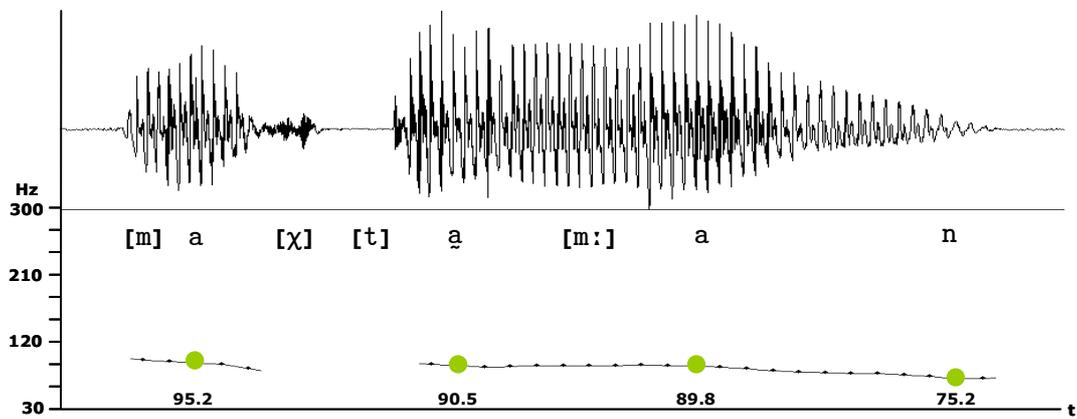


Abb. 20 b

**Sprechfassung III**

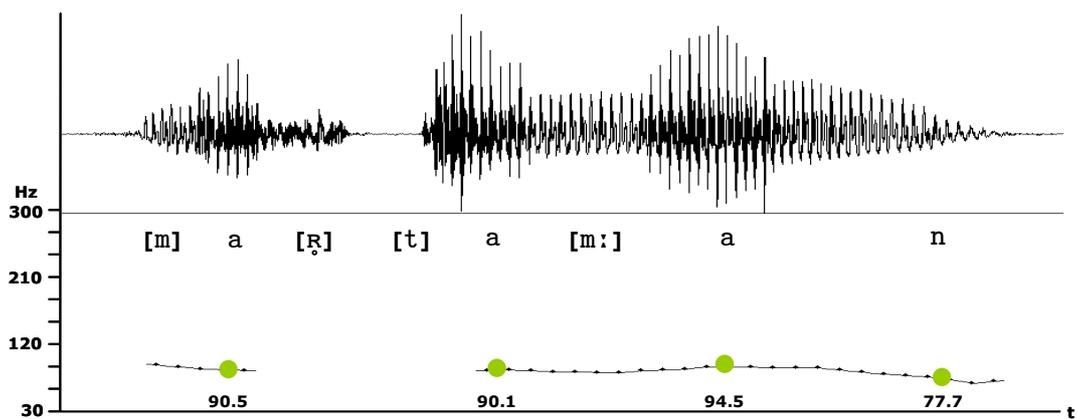


Abb. 20 c

**Abb. 20 a-c: Grundfrequenzverlauf von „Macht am Mann“ aller drei Sprechfassungen im Vergleich. Alle Angaben in Hz.**

der Basstimmlage in die hohe Tenor- bzw. Altstimmlage führt. Damit unterscheidet sich SF I von den anderen beiden SF auch in ihrer Tonmusterrealisation. Während in SF II und III eher terminale Tonmuster gebildet werden, besticht SF I durch eine klare interrogative Tonmusterrealisation (vgl. Abb. 19 und Abb. 20 a-c).

Aus den Messpunkten, die in Abbildung 20 a-c abgetragen sind, wird ebenfalls ersichtlich, dass zentrale Charakteristika der Melodieverläufe in dem alveolaren Nasalkonsonanten manifestiert sind, weswegen es nicht ausreicht, die Messpunkte allein in den Vokalen zu setzen, weil damit wichtige für die Beschreibung melodischer Strukturen notwendige Informationen unberücksichtigt blieben. Im Nachakzentintervall verlaufen SF II und SF III, ähnlich wie bereits in der vorausgegangenen Sequenz, parallel, wobei SF II wiederum einen Halbton unter SF III verläuft. In beiden Sprechfassungen fällt dabei die Melodie vom Akzent ausgehend um jeweils um 3 Halbtonschritte (kleine Terz) ab, wie man aus Abbildung 19 entnehmen kann.

Wenn man die Binnenstruktur der einzelnen SF genauer betrachtet (vgl. Abb. 19), unterscheidet sich auch das Vorakzentintervall, was über die SF hinweg zwar wesentlich homogener gestaltet wird, im Detail. So lässt sich für SF II vom ersten Messpunkt an bis zum Akzent (Messpunkt 3) ein kontinuierliches Abfallen der Melodiekurve in Halbtonschritten beobachten (G-FIS-F), wohingegen sich in SF III die Verlaufskurve bis zum Akzent (Messpunkt 3) ganz monoton auf nur einem Halbton (FIS) bewegt. SF I wiederum ist durch kleinschrittige Auf- und Abbewegungen zwischen den einzelnen Messpunkten (FIS-G-E) markiert.

Eine Übersicht der einzelnen  $F_0$ -Werte und Halbtonschritte ist in Tabelle 16 zusammengestellt. Aus den Intervall-Zeilen und der grau unterlegten Spalte in Tabelle 16 ergibt sich darüber hinaus ein Vergleich zwischen den Halbtonschrittumfängen, die sich aus den Extrempunkten errechnen. Daraus lässt sich ableiten, dass SF I sowohl qualitativ die größten als auch quantitativ die meisten Schwankungen aufweist und damit melodisch am variabelsten erscheint. Im Vergleich dazu erscheint SF II mit 5 Halbtonschritten Gesamtumfang weniger bewegt und SF III mit gerade 3 Halbtonschritten Gesamtumfang nahezu monoton.

Unter Gesichtspunkten der Melodisierung bleibt abschließend noch die Betrachtung der Schlussequenz „das Markante am Mann“. In Abbildung 21 a-c sind die Kurven in ihren Verläufen für die einzelnen SF dargestellt.

### Macht am Mann.

	[m]	a	[χ]	[t]	a	[m:]	a	n	U <sub>I</sub>
<b>SF I</b>		94.9 FIS			95.1 G		84.1 E	218.0 a	
<b>Intervall</b>			1 HTS ↓			3 HTS ↓		17 HTS ↑	17 HTS
	[m]	a	[χ]	[t]	a	[m:]	a	n	U <sub>I</sub>
<b>SF II</b>		95.2 G			90.5 FIS		89.8 F	75.2 D	
<b>Intervall</b>			1 HTS ↓			1 HTS ↓		3 HTS ↓	5 HTS
	[m]	a	[ʀ]	[t]	a	[m:]	a	n	U <sub>I</sub>
<b>SF III</b>		90.5 FIS			90.1 FIS		94.5 FIS	77.7 DIS	
<b>Intervall</b>			0 HTS			0 HTS		3 HTS ↓	3 HTS

**Tab. 16: Melodieverlauf von „Macht am Mann“ aller drei Sprechfassungen im Überblick. Alle Angaben in Hz und Halbtonschritten. Die grau unterlegte Spalte gibt das Intervall zwischen den Extrempunkten an.**

Daraus lässt sich ersehen, dass wiederum die Notwendigkeit besteht, neben den Vokalen zwei zusätzliche Messpunkte auf die sich anschließenden alveolaren Nasalkonsonanten (Messpunkte 4 und 6) zu manifestieren, um die Melodiekurven hinreichend präzise bezüglich ihrer typischen Verläufe beschreiben zu können. Es ergeben sich dabei auf den ersten Blick auch wieder Parallelen zwischen SF II und SF III, wohingegen in SF I größere Frequenzbereiche ausgenutzt werden.

Die Kurven in Abbildung 21 a-c weisen ähnliche Verläufe zu den vorausgegangenen Sprechsequenzen auf. Auch hier hebt sich SF I durch das Nutzen größerer Frequenzbereiche von den anderen beiden SF ab. Vor allem im mittleren Teil (Messpunkte 3 und 4) ist ein starker Melodiegipfel zu verzeichnen, der, wie bereits in der vorherigen Sequenz, sein Maximum erst verspätet durch eine zusätzliche Gleitbewegung vom Vokal in den Nasalkonsonanten erreicht und damit im Vergleich zu SF II und SF III einen zweistufigen Melodieanstieg

**Sprechfassung I**

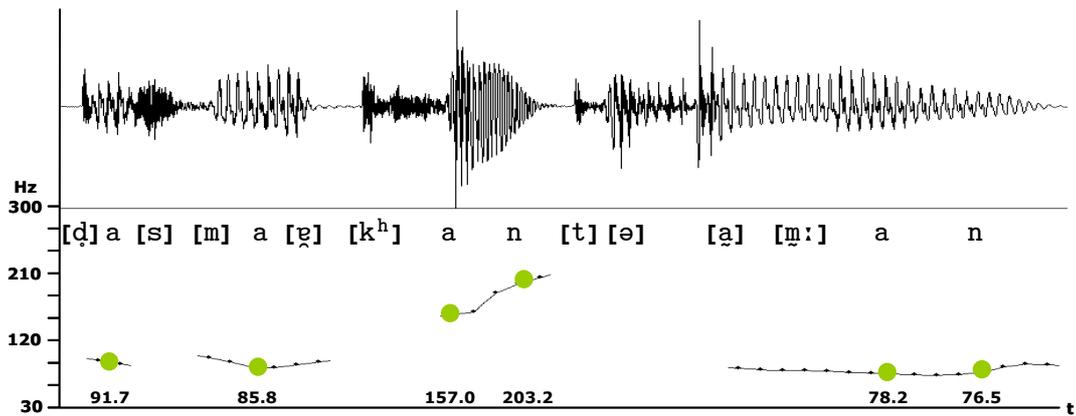


Abb. 21 a

**Sprechfassung II**

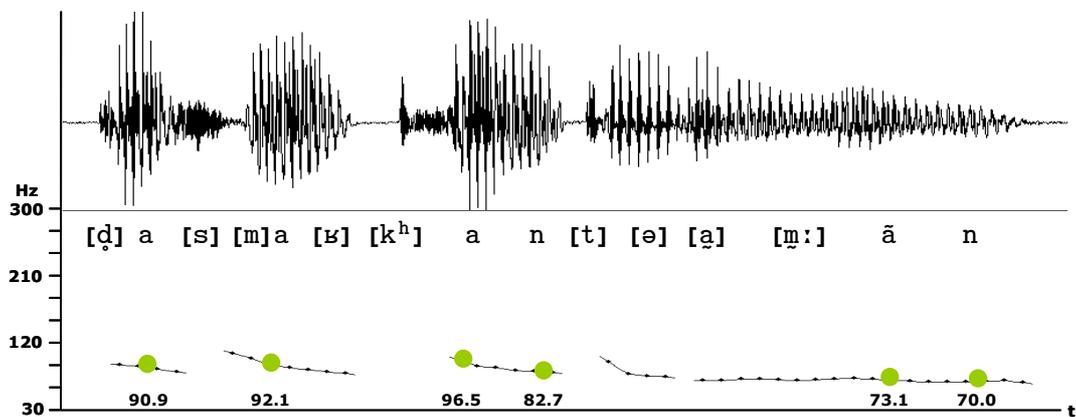


Abb. 21 b

**Sprechfassung III**

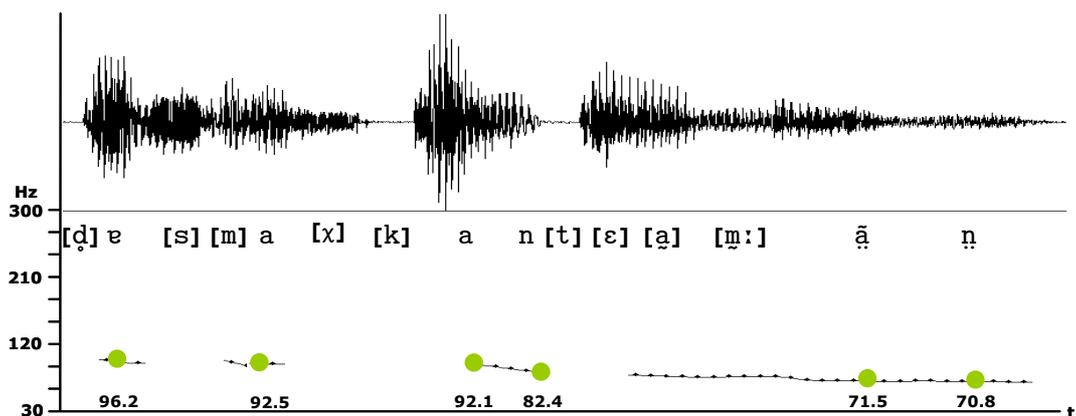
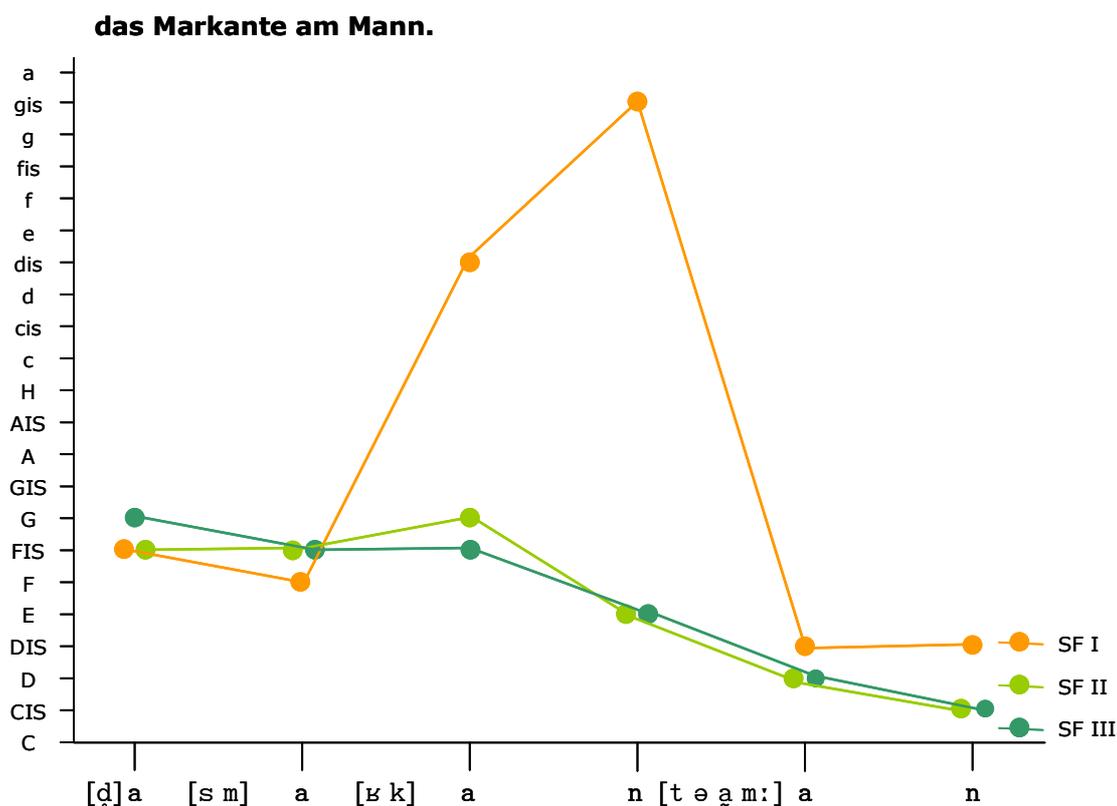


Abb. 21 c

**Abb. 21 a-c: Grundfrequenzverlauf von „das Markante am Mann“ aller drei Sprechfassungen im Vergleich. Alle Angaben in Hz.**

als Akzent erhält. Dieser verschobene zweistufige Gipfel ist besonders deutlich in den Abbildungen 21a und 22 zu erkennen.

Abbildung 22 zeigt wiederum, dass in SF I sowohl quantitativ als auch qualitativ eine Melodiekurve realisiert wird, die sich von den anderen SF hörbar unterscheidet. Mit 17 Halbtönen Differenz zwischen den Messpunkten 4 und 5, wird auf engstem Raum ein Intervallsprung abgefangen, der aus der hohen Tenor bzw. Altstimmlage kommend in die tiefe Bassstimmlage zurückkehrt.



**Abb. 22: Darstellung des melodischen Verlaufs von „das Markante am Mann“ in Halbtonschritten in allen drei Sprechfassungen im Vergleich.**

Dabei wird ebenfalls ein Registerwechsel von der Kopfstimme (Messpunkte 3 und 4) zurück in die Bruststimme kompensiert. Besonders markant treten die Akzente zudem aus ihrem Umfeld heraus, weil das Vorakzentintervall zusätzlich tiefer gelegt wurde. Derartige Bewegungen sind in SF II und SF III nicht zu beobachten. Während das Nachakzentintervall vom Akzent an völlig parallel verläuft, wobei sich die Melodie Halbton um Halbton bis CIS absenkt und damit die gesamte Äußerung in der Lösungstiefe, die sich durch eine tiefe Bass-

stimmlage auszeichnet, beendet, ist im Vorakzentintervall eine gegenläufige Melodiebewegung in SF II und SF III zu beobachten. Hier steigt in SF II die Melodie von FIS kommend bis G um einen Halbtonschritt an, wohingegen in SF III die Melodie von G kommend um einen Halbton zu FIS abfällt. Damit liegen die Akzente, die nur minimal aus ihrem Umfeld heraustreten, um einen Halbton voneinander entfernt. In der folgenden Tabelle sind die  $F_0$ -Verläufe und die entsprechenden Halbtöne für diese Sequenz abschließend vergleichend dargestellt.

### das Markante am Mann.

	[d̥]	a	[s]	[m]	a	[ɐ]	[k <sup>h</sup> ]	a	n	[t]	[ə]	U <sub>I</sub>
SF I		91.7 FIS			85.8 F			157.0 dis	203.2 gis			
Intervall					1 HTS ↓		10 HTS ↑		5 HTS ↑			
	[ạ]	[m:]	a	n								
			78.2 DIS	76.6 DIS								
Intervall		7 HTS ↓		0 HTS								17 HTS
	[d̥]	a	[s]	[m]	a	[ɤ]	[k <sup>h</sup> ]	a	n	[t]	[ə]	U <sub>I</sub>
SF II		90.9 FIS			92.1 FIS			96.5 G	82.7 E			
Intervall					0 HTS		1 HTS ↑		3 HTS ↓			
	[ạ]	[m:]	ā	n								
			73.1 D	70.0 CIS								
Intervall		2 HTS ↓		1 HTS ↓								6 HTS
	[d̥]	e	[s]	[m]	a	[χ]	[k]	a	n	[t]	[ɛ]	U <sub>I</sub>
SF III		96.2 G			92.5 FIS			92.1 FIS	82.4 E			
Intervall					1 HTS ↓				2 HTS ↓			
	[ạ]	[m:]	ã	ṅ								
			71.5 D	70.8 CIS								
Intervall		2 HTS ↓		1 HTS ↓								6 HTS

**Tab. 17: Melodieverlauf von „das Markante am Mann“ aller drei Sprechfassungen im Überblick. Alle Angaben in Hz und Halbtonschritten. Die grau unterlegte Spalte gibt das Intervall zwischen den Extrempunkten an.**

Aus Tabelle 17 kann man entnehmen, dass auch in der Abschlussequenz in SF I sowohl unter qualitativen Aspekten größere melodische Intervalle, als

auch unter quantitativen Gesichtspunkten mehr melodische Veränderungen realisiert werden und sich SF I damit in ihrer melodischen Variabilität deutlich von SF II und SF III absetzt. Mit 6 Halbtonschritten Differenz zwischen den Extrempunkten in SF II und SF III wird im Vergleich zu 17 Halbtonschritten Differenz in SF I gerade ein Drittel an Spielraum genutzt.

Dieses extrem große Intervall von über einer großen Dezime kann nur unter Ausnutzung der Kopfstimme entstehen, was bedeutet, dass die melodischen Akzente in SF I aus der Modalstimme in die Kopfstimme verlegt sind, was ein weiteres differierendes Merkmal gegenüber SF II und SF III darstellt.

Zusammenfassend kann festgehalten werden, dass die Betrachtung melodischer Veränderlichkeiten innerhalb der einzelnen SF unter phonetischen Gesichtspunkten sowohl unter qualitativen, als auch unter quantitativen Gesichtspunkten zu einer Diskriminierung der einzelnen SF voneinander führt. Während sich SF I in allen Teilsequenzen durch größere Intervalle und häufigere Melodiebewegungen deutlich von SF II und SF III absetzt, unterscheiden sich diese beiden weniger stark in ihrer melodischen Variabilität voneinander. Obwohl beide Fassungen durchgängig in der Bruststimme realisiert sind, liegt SF II in der Tendenz allerdings stets unter SF III. Inwieweit diese Unterschiede auch hörbar sind, soll im folgenden Kapitel beschrieben werden.

## 6.2 AUDITIVE ANALYSE

Im vorangegangenen Kapitel wurden Differenzen der einzelnen Sprechweisen unter phonetischen Gesichtspunkten der Temporalität, der Artikulation und der Melodisierung beschrieben, wie sie im akustischen Signal messbar sind. Dabei handelt es sich in einigen Fällen um sehr kleine Unterschiede. Es stellt sich von daher die berechtigte Frage, ob und inwieweit diese Unterschiede innerhalb der einzelnen SF auf Grund der Kürze des Sprechanteils überhaupt wahrnehmbar sind. Zudem korrelieren die in der Akustik verwendeten Skalen nicht immer linear mit dem menschlichen Gehör (vgl. Kap. 4.1), so dass in der vorliegenden Untersuchung zusätzlich zu der phonetischen Analyse eine auditive

Analyse in Form eines Expertenurteils<sup>13</sup> erfolgte, deren Ergebnisse in diesem Kapitel genauer beschrieben werden sollen.

Wie im Methodenteil (Kap. 5.2) ausgeführt, dienten die Merkmalskataloge von Fährmann (1967) und Kranich (2003) als Vorlage für die Auswahl der Kriterien für die auditive Beurteilung der SF. In einem Fragebogen (siehe Anhang) wurden dabei (a) die habituellen Stimmqualitätsmerkmale: Sprechstimmlage (tief, mittel, hoch), Intensität (gering, mittel, hoch), Sprechgeschwindigkeit (langsam, mittel, schnell), Stimmeinsätze und Stimmqualität (klar, verhaucht, knarrend), Klangfülle (resonanzreich, mittel, resonanzarm) und Klangfarbe (hell vs. dunkel, weich vs. hart, warm vs. kalt) und (b) die individuellen Verlaufskriterien: Melodische Veränderlichkeit und Dehnungsakzente, erhoben.

Die Auswahl dieser Merkmale erfolgte unter dem Gesichtspunkt, die drei Sprechfassungen hinreichend voneinander diskriminieren zu können. Da es sich bei den Sprechfassungen um sehr kurze Äußerungssequenzen handelt, erwiesen sich einige Merkmale als wenig brauchbar, weswegen sie in der auditiven Analyse keine Berücksichtigung finden.

In Abbildung 23 sind die einzelnen SF in Form von Profilen über die verschiedenen Kriterien der Stimmqualitätsmerkmale hinweg abgetragen, so dass sich ein guter Überblick über die Verschiedenartigkeit in den drei Realisierungen gewinnen lässt.

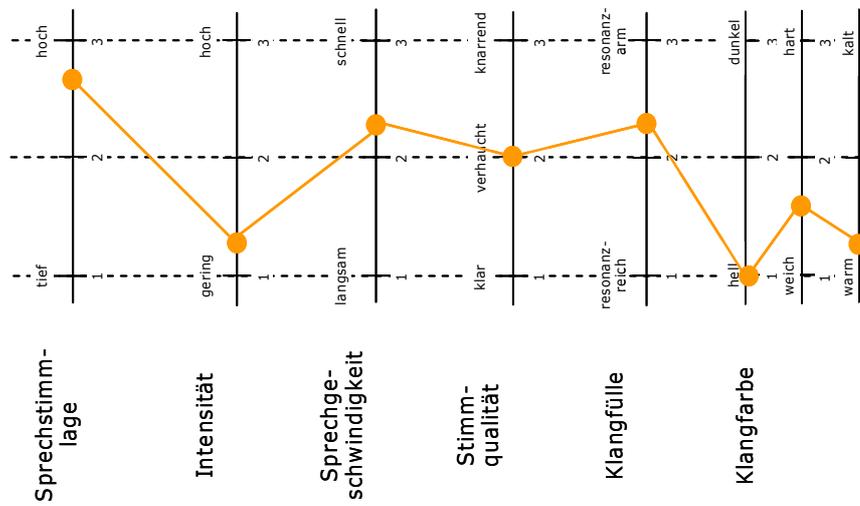
Jedes Stimmmerkmal sollte im Fragebogen durch die Experten auf einer Dreier-Skala eingeschätzt werden. Dabei liegen den Eigenschaften in allen Fällen ordinalskalierte Merkmale zu Grunde, die einzige Ausnahme bildet der Punkt der Stimmqualität, hierbei handelt es sich um eine nominale Skalierung. Die in Abbildung 23 dargestellten Punkte entsprechen den gebildeten Mittelwerten (MW), die sich aus den sechs Hörurteilen ergeben. Aus den Verbindungslinien der einzelnen Punkte entstehen für die SF typische Profile, die bereits beim flüchtigen Überblicken auf hörbare Unterschiede zwischen den Stimmqualitätsmerkmalen hinweisen, und die im Folgenden etwas detaillierter beschrieben werden sollen.

Bei der Einschätzung der Sprechstimmlage (tief = 1, hoch = 3) unterscheiden sich vor allem SF I und SF III voneinander, die mit Mittelwerten von 2.83 ( $\pm 0.41$ ) vs. 1.00 ( $\pm 0.00$ ) die Extrempositionen einnehmen.

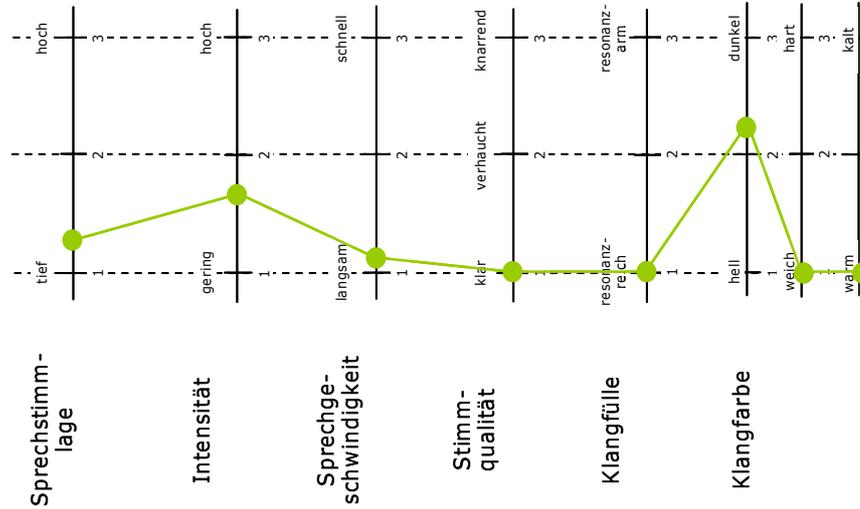
---

<sup>13</sup> Bei der Expertengruppe handelt es sich um sechs ausgebildete und praktizierende Sprechwissenschaftler, deren Professionalität maßgeblich in der auditiven Beurteilung von Stimmen und Sprechleistungen liegt.

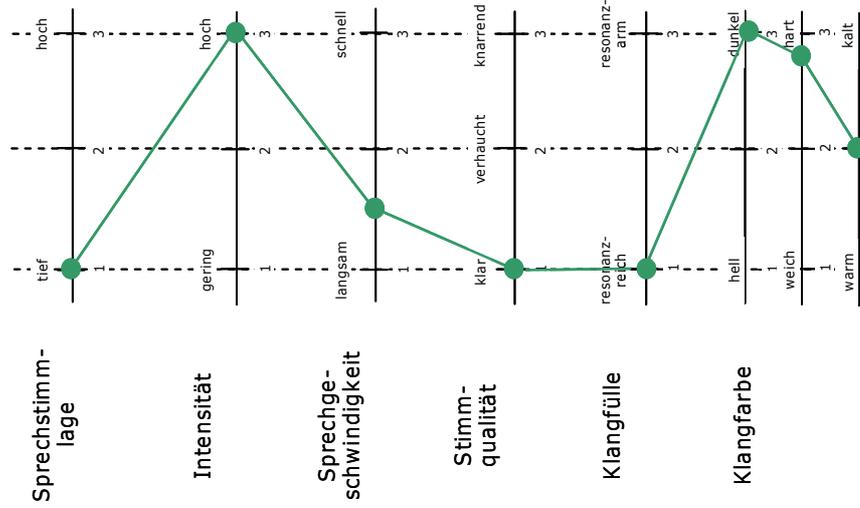
**Profil Sprechfassung I**



**Profil Sprechfassung II**



**Profil Sprechfassung III**



**Abb. 23: Profildarstellung der Stimmqualitätsmerkmale aller drei Sprechfassungen im Vergleich.**

SF II befindet sich mit einem MW von 1.33 ( $\pm 0.52$ ) allerdings in deutlicher Nähe zu SF III.

Diese Einschätzung deckt sich mit den Ergebnissen der apparativen Analyse (vgl. Kap. 6.1), auch hier setzt sich SF I mit höheren  $F_0$ -Werten von SF II und SF III deutlich ab, wobei SF II allerdings tendenziell etwas tiefere Messwerte hervorbrachte als SF III. Diese Differenzen sind aber sowohl bei der apparativen als auch bei der auditiven Analyse so gering, dass die beiden SF in Bezug auf die Beschreibung der Sprechstimmlage als äquivalent betrachtet werden können.

Im Gegensatz zur Sprechstimmlage hebt sich unter Gesichtspunkten der Intensität (gering = 1, hoch = 3) SF III (MW 3.00  $\pm 0.00$ ) deutlich von SF I (MW 1.33  $\pm 0.52$ ) und SF II (MW 1.67  $\pm 0.52$ ) ab. Da sich die einzelnen SF akustisch allerdings nicht bezüglich der Lautstärke unterscheiden, liegt die Vermutung nahe, dass der Terminus Intensität von den Kontrollhörern einheitlich mit Spannungszuständen sowohl an, als auch oberhalb der Glottis in Verbindung gebracht wurde. Die hohe Spannung von SF III zeigt sich im akustischen Signal vor allem im artikulatorischen Bereich, wo starke temporale Markierungen die Folge von gespannten und gedehnten lautlichen Realisationen sind, die sich durch verstärkte Friktionen, fehlenden Aspiration und ausgeprägte Glottalisierungstendenzen auszeichnen (vgl. Kap. 6.1).

Ein weiteres Kriterium stellt die Beurteilung der Sprechgeschwindigkeit (langsam = 1, schnell = 3) dar. Hierbei unterscheiden sich alle drei SF voneinander. Im Durchschnitt wird dabei SF I (MW 2.33  $\pm 0.52$ ) als schneller im Vergleich zu SF II (MW 1.17  $\pm 0.41$ ) und SF III (MW 1.50  $\pm 0.55$ ) eingeschätzt. Diese Rangfolge deckt sich allerdings nicht einheitlich mit den Ergebnissen der apparativen Analyse (vgl. Kap. 6.1). Hier wird zwar ebenfalls SF I mit der kürzesten Realisationszeit ermittelt, aber SF III erstreckt sich über einen größeren zeitlichen Rahmen als SF II. Worauf diese Differenz zurückzuführen ist, wird in der Diskussion der Ergebnisse (Kap. 7) genauer beschrieben. Es soll hier nur die folgende Beobachtung angeführt werden: Betrachtet man ausschließlich die erste Sprechsequenz „Caractère“ und die sich daran anschließende Pause, so ist diese Sequenz für SF II länger als für SF III. Es liegt von daher die Vermutung nahe, dass die Kontrollhörer in ihrem perzeptiven Urteil in Bezug auf die Sprechgeschwindigkeit von Merkmalen der Pausensetzung geleitet

wurden. Untersuchungen, die diesen Befund belegen, hat Goldman-Eisler (1968) durchgeführt (vgl. Kap. 7).

Betrachtet man die unterschiedlichen Profile in Bezug auf die Stimmqualität, so lässt sich erkennen, dass sich SF I mit der Attributierung „verhaucht“ von SF II und SF III abhebt, deren Stimmqualitäten als „klar“ eingeschätzt werden. Dieses Urteil, welches sehr einheitlich ( $\pm 0.00$ ) gefällt wurde, bestätigt die Angaben der apparativen Analyse (vgl. Kap. 6.1). Auch hier lassen sich für SF I hohe Hauchanteile im Spektrogramm nachweisen, wohingegen sich vor allem SF II durch eine hohe Stimmhaftigkeit auszeichnet. Die Beurteilung einer klaren Stimmqualität in SF II und SF III schließt allerdings nicht aus, dass aus Funktionsgründen heraus auch Hauch (für die Aspiration) und Knarren (für den festen Stimmeinsatz) im Signal nachzuweisen sind, die in diesem Fall allerdings nicht als Wirkungsmittel, und damit auch nicht als Eigenschaften der Stimme eingesetzt werden.

Auch unter Gesichtspunkten der Klangfülle (resonanzreich = 1, resonanzarm = 3) hebt sich SF I von den anderen beiden SF ab. So wird hier mit einem MW von 2.33 ( $\pm 0.52$ ) die Stimme als wesentlich resonanzärmer als in SF II und SF III, wo sie mit einem MW von je 1.00 ( $\pm 0.00$ ) als sehr resonanzreich eingeschätzt wird, beurteilt. Da es sich in allen SF um denselben Sprecher handelt, lässt sich zu Recht die Frage danach stellen, wieso sich die Wahrnehmung der Klangfülle der Stimme verändert. Die Erklärung dafür liegt in der Realisierung der Akzente. In SF I werden neben quantitativ vielen melodischen Veränderungen auch qualitativ große melodische Intervalle realisiert. Um diese Variabilität bei kurzer temporaler Ausdehnung zu erreichen, wechselt der Sprecher sehr häufig zwischen Brust- und Kopfreger. Die im Kopfreger veränderte Stimmqualität ist wiederum auf veränderte Spannungszustände an den Stimmlippen zurückzuführen. Im Vergleich zur Bruststimme schwingen hier nicht mehr die ganzen Stimmlippen, sondern nur noch deren Randkanten, wodurch ein schlankerer Ton, der demzufolge auch als deutlich resonanzärmer eingeschätzt wird, entsteht. Wirth (1995:128f) definiert die unterschiedlichen Register dabei folgendermaßen:

Brustregister = Aktivspannung der Stimmlippen (Kontraktion des M. vocalis) = Vollschiwingung. (Wirth 1995:128)

Kopfreger = Passivschwiwingung der Stimmlippen (Kontraktion des M. cricothyroideus) = Randschiwingung. (Wirth 1995:129)

Diese Registerwechsel sind in SF II und SF III nicht eingesetzt. Über die Äußerungssequenz hinweg wird hier ausschließlich das Brustregister verwendet, und somit schwingen auch die gesamten Stimmlippen über den ganzen Äußerungszeitraum, was wiederum die als sehr resonanzreich wahrgenommene Stimmqualität nach sich zieht.

Für die Beurteilung der stimmlichen Merkmale bleibt abschließend noch die Betrachtung der Klangfarbe der Stimme. Hierbei wurden den Kontrollhörern drei Gegensatzpaare, die als Metaphern anderer Sinnesbereiche auf die Stimmqualität übertragen wurden, angeboten (hell vs. dunkel, weich vs. hart und warm vs. kalt). Den Profilen in Abbildung 23 lässt sich entnehmen, dass auch bezüglich des Timbres Unterschiede in den einzelnen SF zu verzeichnen sind. Auf welche Modifikationen an der Glottis selbst oder oberhalb der Glottis im Ansatzrohr die verschiedenen Klangfarben zurückzuführen sind, und wie sie sich im akustischen Signal darstellen lassen, ist bis heute noch nicht zufrieden stellend beantwortet (vgl. Kap. 4.1). Von daher seien nur die auditiv ermittelten Ergebnisse der Kontrollhörer vorgestellt.

Im ersten Paar (hell = 1, dunkel = 3) differieren die Daten von SF I (MW 1.00  $\pm$  0.00) über SF II (MW 2.20  $\pm$  0.45) zu SF III (MW 3.00  $\pm$  0.00) und umfassen damit die gesamte Skalenbreite. Diese Befunde lassen sich noch am ehesten mit den zuvor beschriebenen Resonanzverhältnissen erklären. Die häufige Verwendung des Kopffregisters in SF I und der daraus resultierende schlanke Ton lässt das Timbre der Stimme dementsprechend hell erscheinen, wohingegen die beschriebene Rückverlagerungstendenz in SF III eher eine dunkle Klangfarbe der Stimme mit sich bringt. Beide Phänomene sind in SF II nicht zu beobachten, von daher wird ihr weder eine helle, noch eine besonders dunkel anmutende Klangfarbe zugesprochen.

Ein zweites Paar stellt die Opposition der Eigenschaften weich (= 1) vs. hart (= 3) dar. Auch hier ist nahezu die gesamte Skalenbreite ausgeschöpft. So wird SF II mit einem MW von 1.00 ( $\pm$  0.00) als ausnahmslos weich empfunden, wohingegen sich SF I mit einem MW von 1.67 ( $\pm$  0.52) leicht zentralisiert und SF III mit einem MW von 2.83 ( $\pm$  0.41) nahezu den anderen Pol, der als hart empfunden wird, erreicht.

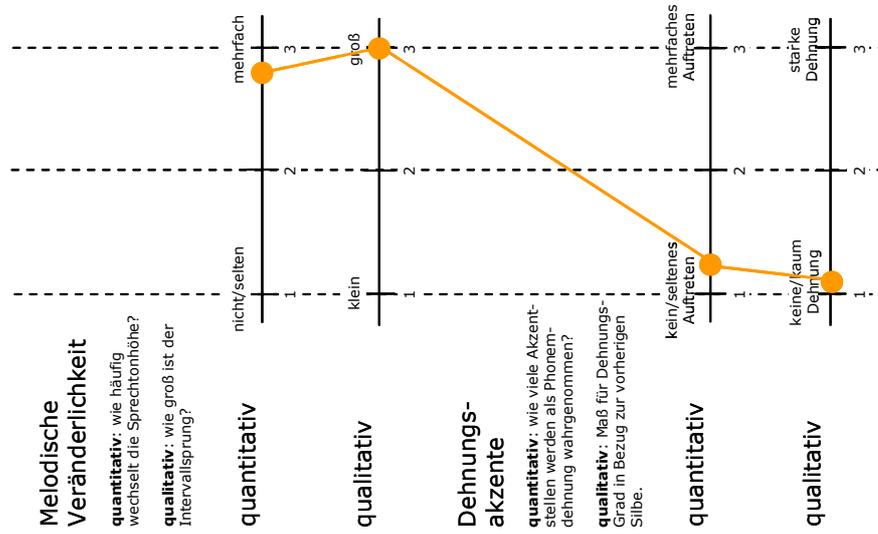
Weniger intensiv werden die Differenzen bezüglich des dritten Paares warm (= 1) vs. kalt (= 3) empfunden. Mit MW von 1.33 ( $\pm$  0.52) für SF I, 1.00 ( $\pm$  0.00) für SF II und 2.00 ( $\pm$  0.00) für SF III liegen die Höreindrücke deutlich enger

beieinander als in den anderen Eigenschaftspaaren. Es lässt sich die Vermutung aufstellen, dass die Opposition von warm vs. kalt in diesem Kontext separat betrachtet weniger aussagekräftig ist und von daher vielmehr in Beziehung zu anderen Oppositionen stehend wie z.B. weich und hart, betrachtet werden muss. Wenn man in Abbildung 23 diese Verbindung betrachtet, so lässt sich erkennen, dass sich SF II durch die Besetzung der vorderen Extrempositionen (weich und warm) von SF I, wo die MW zwar noch im vorderen Drittel, aber zentralisiert liegen und SF III wo die MW deutlich in Richtung der hinteren Extrempositionen (hart und kalt) verschoben liegen, abhebt.

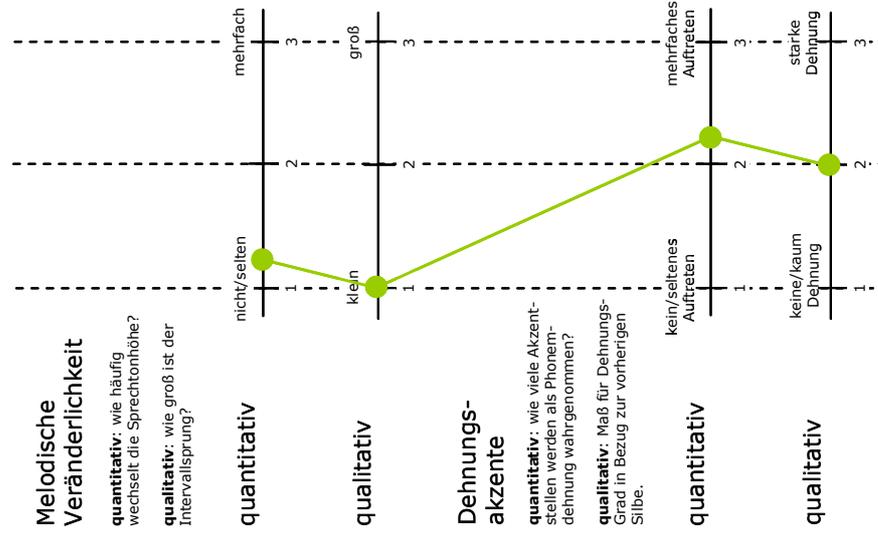
Aus den Ausführungen zeigt sich, dass die Profile der einzelnen SF bezüglich der Stimmqualitätsmerkmale nachweisbar divergieren. Es sollen nun die Ergebnisse der Verlaufsmerkmale, die die sprecherische Realisierung der Äußerung beschreiben, aufgezeigt werden. In Abbildung 24 ist auch dafür eine Profilkurve der einzelnen SF erstellt worden, die die zwei Verlaufskriterien der melodischen Veränderlichkeit und der Dehnungsakzente jeweils unter quantitativen als auch unter qualitativen Aspekten beschreibt. Die quantitativen Veränderungen beschreiben dabei wie häufig die Sprechtonhöhe zwischen nicht/selten und mehrfach wechselt bzw. wie viele Akzentstellen als Dehnungen wahrgenommen werden. Die qualitativen Ergebnisse geben Auskunft über die Größe der Intervallsprünge bzw. den Grad der Dehnung im Vergleich zu deren Umfeld (vgl. Kap. 4.1).

Aus den MW, die in Abbildung 24 dargestellt sind, lassen sich zwei Tendenzen ableiten: (a) sowohl in Bezug auf die melodische Veränderlichkeit als auch auf die Dehnungsakzente wird über die einzelnen SF hinweg der gesamte Skalenbereich ausgenutzt und (b) die MW der quantitativ und qualitativ bestimmten Ergebnisse liegen für die zwei jeweiligen Kategorien auffällig dicht beieinander. Im Detail betrachtet zeigen die MW für die melodische Veränderlichkeit in SF I mit 2.83 ( $\pm 0.41$ ) für die quantitative und 3.00 ( $\pm 0.00$ ) für die qualitative Variabilität die höchsten Werte. Das gegenteilige Extrem findet sich in SF II, wo die MW mit 1.33 ( $\pm 0.82$ ) für die quantitative und 1.00 ( $\pm 0.00$ ) für die qualitative Beurteilung die niedrigsten Werte ergeben. Zwischen beiden Polen ordnet sich SF III ein, die mit MW von 2.00 ( $\pm 0.89$ ) und 1.83 ( $\pm 0.75$ ) für die quantitative und für die qualitative Einschätzung der melodischen Veränderlichkeit die Mittelposition einnimmt. Diese Werte decken sich grundsätzlich mit den Ergebnissen der apparativen Untersuchung (vgl. Kap. 6.1). Hier konnte sehr

### Profil Sprechfassung I



### Profil Sprechfassung II



### Profil Sprechfassung III

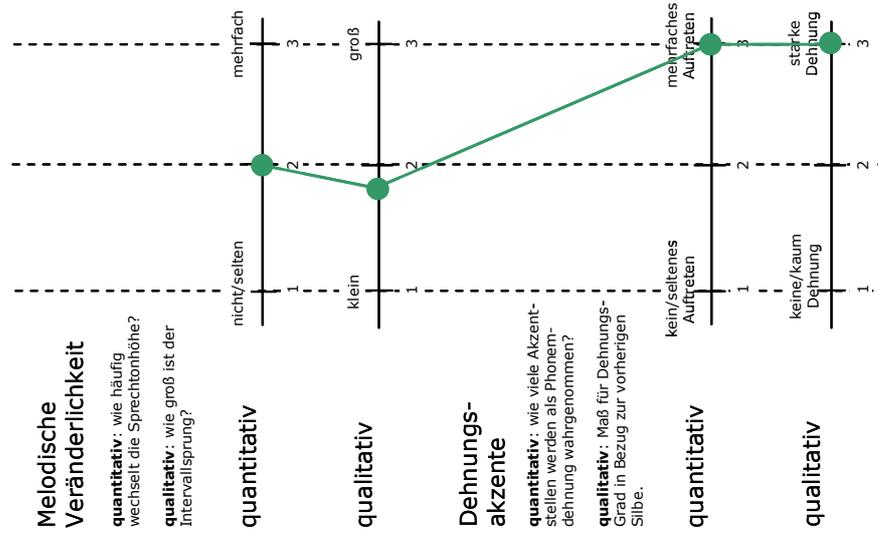


Abb. 24: Profildarstellung der Verlaufskriterien aller drei Sprechfassungen im Vergleich.

klar aufgezeigt werden, das sich SF I mit deutlich erhöhten Werten für den  $F_0$ -Verlauf sowohl quantitativ als auch qualitativ von SF II und SF III absetzt, die im akustischen Signal unter Gesichtspunkten der Melodisierung im Vergleich zur auditiven Analyse allerdings relativ eng beieinander liegen.

Betrachtet man den Profilverlauf vom Übergang aus der melodischen Veränderlichkeit zur Einschätzung der Dehnungsakzente, so lässt sich SF I mit einem diagonalen Wechsel über zwei Skalenstufen (von stark melodisch variiert zu keine Verwendung von Dehnungsakzenten) von SF II und SF III deutlich abheben, die jeweils nur eine Stufe auf der Skala zwischen den beurteilten Kriterien überwinden. Mit MW von  $1.33 (\pm 0.52)$  und  $1.20 (\pm 0.45)$  liegen die Angaben für die Quantität und die Qualität der Dehnungsakzente in SF I den Werten der Melodisierung diametral entgegen. Das Ausnutzen der Extrembereiche auf der Skala ist für SF I typisch und unterscheidet sie von SF II und SF III hörbar.

Was die Akzentsetzung durch Dehnung betrifft nimmt SF II mit MW von  $2.33 (\pm 0.52)$  und  $2.00 (\pm 0.00)$  unter quantitativen und qualitativen Aspekten eine mittlere Position ein, die von SF III mit mehrfach wahrgenommenen Dehnungen ( $MW\ 3.00 \pm 0.00$ ) und zusätzlich starker Ausprägung ( $MW\ 3.00 \pm 0.00$ ) hörbar übertroffen wird. Zu dieser Einschätzung kamen die Kontrollhörer sehr einheitlich. Die geschilderten Höreindrücke lassen sich auch durch die Messungen im akustischen Signal verifizieren (vgl. Kap. 6.1), wo sich SF III ebenfalls durch die Dehnung einzelner Lautsegmente deutlich von SF I und SF II absetzt, und dadurch auch ihre charakteristische Ausdrucksform erhält.

In Tabelle 18 sind abschließend noch einmal alle Mittelwerte (MW) und Standardabweichungen (STAW) für die Stimmqualitätsmerkmale und die Verlaufsmerkmale überblicksartig zusammengetragen. Daraus lässt sich zusammenfassend festhalten, dass die im akustischen Signal untersuchten Kriterien einen Niederschlag in der auditiven Analyse finden und damit auch perzipiert werden.

In den Grundaussagen decken sich dabei die Ergebnisse der apparativen Analyse mit den Ergebnissen des Expertenurteils, so dass sich klare Tendenzen für die einzelnen SF aufzeigen lassen, die sie voneinander diskriminieren. An einigen Stellen konnte demonstriert werden, dass beide Untersuchungsmethoden zu genau denselben Ergebnissen führen (z.B. Stimmqualität, Dehnungsakzente), an anderen Stellen zeigten sich kleinere Abweichung in der Rangfolge der

einzelnen SF (z.B. Sprechgeschwindigkeit, melodische Veränderlichkeit), die aber in keinem der Fälle grundlegenden Tendenzen widersprechen.

So kann resümiert werden, dass, obwohl es sich teilweise um nur sehr kleine Nuancen an Variation einzelner Parameter zwischen den SF handelt, diese von aufmerksamen und geschulten Hörern überindividuell zuverlässig erkannt und zugeordnet werden können.

### Übersicht Daten Expertenurteil

	Fassung I		Fassung II		Fassung III	
	MW	STAW	MW	STAW	MW	STAW
<b>Sprechstimmlage</b> (tief = 1, hoch = 3)	2.83	± 0.41	1.33	± 0.52	1.00	± 0.00
<b>Intensität</b> (gering = 1, hoch = 3)	1.33	± 0.52	1.67	± 0.52	3.00	± 0.00
<b>Sprechgeschwindigkeit</b> (langsam = 1, schnell = 3)	2.33	± 0.52	1.17	± 0.41	1.50	± 0.55
<b>Stimmqualität</b> (klar = 1, verhaucht = 2, knarrend = 3)	2.00	± 0.00	1.00	± 0.00	1.00	± 0.00
<b>Klangfülle</b> (resonanzreich = 1, resonanzarm = 3)	2.33	± 0.52	1.00	± 0.00	1.00	± 0.00
<b>Klangfarbe</b>						
<b>hell</b> (= 1) – <b>dunkel</b> (= 3)	1.00	± 0.00	2.20	± 0.45	3.00	± 0.00
<b>weich</b> (= 1) – <b>hart</b> (= 3)	1.67	± 0.52	1.00	± 0.00	2.83	± 0.41
<b>warm</b> (= 1) – <b>kalt</b> (= 3)	1.33	± 0.52	1.00	± 0.00	2.00	± 0.00
<b>Melodische Veränderlichkeit</b>						
<b>quantitativ</b> (nicht/selten = 1, mehrfach = 3)	2.83	± 0.41	1.33	± 0.82	2.00	± 0.89
<b>qualitativ</b> (klein = 1, groß = 3)	3.00	± 0.00	1.00	± 0.00	1.83	± 0.75
<b>Dehnungsakzente</b>						
<b>quantitativ</b> (kein/selten = 1, mehrfach = 3)	1.33	± 0.52	2.33	± 0.52	3.00	± 0.00
<b>qualitativ</b> (kein/kaum = 1, stark = 3)	1.20	± 0.45	2.00	± 0.00	3.00	± 0.00

**Tab. 18: Übersicht über die Mittelwerte (MW) und Standardabweichungen (STAW) der auditiven Analyse (gemittelte Expertenurteile).**

Als Zwischenergebnis der phonetischen Analyse, die sich aus der apparativen und der auditiven Untersuchung zusammensetzt, kann eine ganz klare Tendenz dahingehend aufgezeigt werden, dass SF II über die verschiedenen Kriterien hinweg niemals eine Extremposition einnimmt. Dieser Befund spricht da-

für, dass sich SF II durch ein ausgewogenes Verhältnis an stimmlicher und sprecherischer Modulation von SF I (die unter melodischen Gesichtspunkten durch extreme Variabilität heraus sticht) und SF III (die unter temporalen und artikulatorischen Gesichtspunkten durch extreme Variabilität heraus sticht) durch ein gleichmäßiges Zusammenwirken verschiedener prosodischer Merkmale deutlich abhebt.

Ob nun nicht-geschulte Hörer und Laien auf diese stimmlichen und sprecherischen, teilweise sehr feinen Unterschiede reagieren, ohne dafür vorab sensibilisiert worden zu sein (Blindstudie), soll die sich anschließende Fragebogenuntersuchung zeigen, die am Beispiel der oben besprochenen Parfumwerbung durchgeführt wurde. Darüber hinaus soll der Frage nachgegangen werden, ob es eine SF gibt, die generelle Akzeptanz findet und damit den Werbespot in seiner Wirkung positiv beeinflusst, oder ob alle drei SF in Bezug auf bestimmte Wirkungskriterien gleichermaßen akzeptiert werden.

### 6.3 WIRKUNGSANALYSE

„Durchs Auge lieben?  
Nichts ist abgeschmackter.  
Der Kehlkopf nur verrät uns den Charakter.“  
Wilhelm Jordan<sup>14</sup>

Nachdem die drei SF bezüglich ihrer Differenzen unter phonetischen Gesichtspunkten sowohl apparativ als auch auditiv beschrieben wurden, soll nun aufgezeigt werden, inwieweit diese Unterschiede auch zu einer möglichen Änderung in der Wahrnehmung des Werbespots führen. Dazu wurde einer zufällig gewählten Stichprobe von 156 Studenten verschiedener Fakultäten (naturwissenschaftliche, sozialwissenschaftliche und geisteswissenschaftliche Fakultät) der Friedrich-Schiller-Universität Jena das oben besprochene Werbevideo vorgespielt, welches die Probanden dann in Form eines Fragebogens (siehe Anhang) mit Aussagen zu verschiedenen Attributen des Werbespots, der dargestellten Person und des erwarteten Duftes beurteilen sollten (vgl. Kap. 5.2).

---

<sup>14</sup> aus dem Lustspiel „Durchs Ohr“.

Da alle Probanden zum Zeitpunkt der Umfrage Studierende waren, kann die Stichprobe bezüglich des Alters als ausgesprochen homogen beschrieben werden. In Abbildung 25 a und b sind die Alters- und die Geschlechterverteilung graphisch abgebildet. Mit 22.1% und 61.0% nehmen die Gruppen der unter 20 Jährigen, bzw. der 21-25 Jährigen mehr als drei Viertel und damit den größten Teil der Stichprobe ein. Die Gruppe der 25-28 Jährigen (10.4%) und die Gruppe <29 (mit lediglich 6.5%) stellen nur einen kleineren Teil der Probanden.

Abb. 25 a

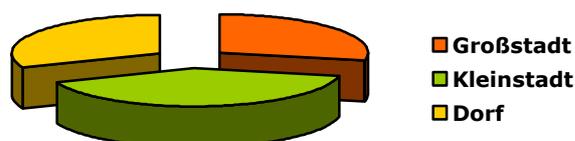
Abb. 25 b



**Abb. 25 a und b: Graphische Darstellung der Altersverteilung und der Geschlechterverteilung innerhalb der Stichprobe.**

Unter Gesichtspunkten der Geschlechterverteilung sind die Frauen mit 82.1% die stärkste Gruppe. Die männlichen Probanden nehmen mit 19.9% circa ein Fünftel der Stichprobe ein.

Für die Beschreibung der Stichprobe wurde ebenfalls erhoben, ob die Probanden in einer Großstadt (< 100.000 Einwohner), in einer Kleinstadt (5.000 – 100.000 Einwohner) oder in einem Dorf ( $\leq$  5.000 Einwohner) aufgewachsen sind. Die Ergebnisse, die in Abbildung 26 zusammengefasst sind, zeigen folgende Aufteilung: 40.6% aller Probanden der Stichprobe leben in einer Kleinstadt, 31.0% in einem Dorf und 28.4% in einer Großstadt. Wie man in der Graphik erkennen kann, sind alle drei Kriterien, die der Beschreibung der regionalen Herkunft dienen, in einem vergleichbaren Anteil in der Stichprobe repräsentiert.



**Abb. 26: Zusammensetzung der Stichprobe bezüglich der regionalen Herkunft nach Großstadt, Kleinstadt und Dorf.**

Die Stichprobe setzte sich für die Untersuchung aus drei Gruppen zusammen, wobei jede Gruppe nur eine der oben beschriebenen SF vorgestellt bekam und beurteilen sollte. Dazu wurde das Video in jeder Gruppe drei Mal hintereinanderweg vorgespielt. Das Untersuchungsdesign war von der Frage nach der Perzeption prosodischer Merkmale im Persuasionsprozess einer Parfumwerbung geleitet, obwohl die Probanden vorab in keiner Weise für stimmliche und sprecherische Merkmale des Werbespots sensibilisiert wurden, um einen systematischen Fehler durch eine gelenkte Aufmerksamkeit der Versuchspersonen auszuschließen (Blindstudie). Ziel des Fragebogens war es, herauszufinden, ob die Veränderung stimmlicher und sprecherischer Merkmale verschiedener Sprechweisen eines Werbeslogans zu einer veränderten Wahrnehmung und damit auch Wirkung eines Werbespots führt.

In einem ersten Abschnitt des Fragebogens bekamen die Probanden die Eigenschaften exklusiv, edel, farblos, professionell, schlicht, preiswert, animierend, gewöhnlich vorgelegt, die einen Werbespot näher spezifizieren. Diese Items sollten jeweils auf einer 5-er Skala von -2 (trifft gar nicht zu) bis +2 (trifft völlig zu) anhand des vorgeführten Werbevideos beurteilt werden. Die Auswahl der Eigenschaften war durch die Zuordnung von positiven und negativen Items geleitet. Um sicher zu stellen, dass die Probanden keine willkürlichen Kreuze setzten, sondern die einzelnen Items systematisch bewerteten, wurde in einem ersten Schritt durch eine Faktorenanalyse untersucht, inwieweit die genannten Items in einem Set von Variablen miteinander korrelieren, d.h. inwieweit den einzelnen Variablen (Konstrukten) eine gemeinsame bzw. übergeordnete Idee zugrunde liegt. Die Faktorenanalyse ergab für die oben genannten Items des Werbespots eine klare Zuordnung in eine Gruppe positiver Merkmale (exklusiv, edel, professionell, animierend) und eine Gruppe negativer Merkmale (schlicht, preiswert, gewöhnlich), wodurch die intendierte Einteilung durch die Faktorenanalyse bestätigt wurde. Das Merkmal farblos fiel aus beiden Faktoren (Werbespot<sub>positiv</sub> und Werbespot<sub>negativ</sub>) heraus, was möglicherweise auf die schwarz-weiss-Darstellung des Videos zurückzuführen ist. Aus diesem Grund wird das Item farblos aus der folgenden Analyse ausgeschlossen und findet im Weiteren keine Berücksichtigung mehr.

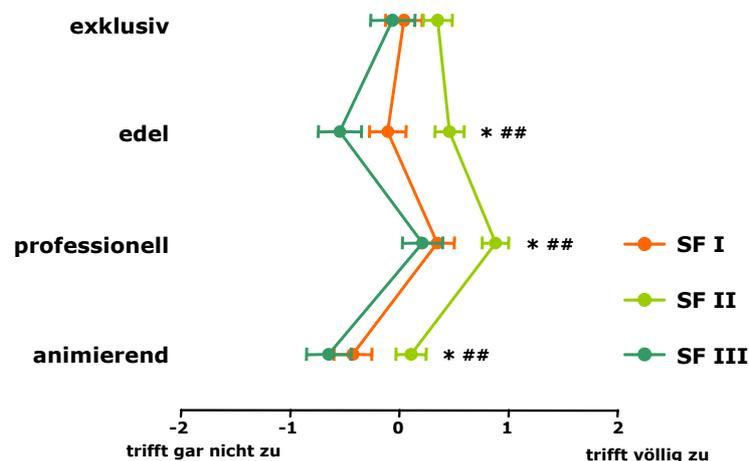
In den folgenden Abbildungen (Abb. 27-34) werden jeweils die einzelnen Items der untersuchten Faktoren für alle drei SF profilartig im Vergleich dargestellt. Diese Form der Darstellung ermöglicht eine systematische Präsentation

der Daten, aus der sich schnell die Bewertungen der einzelnen SF ablesen lassen.

### Perzeption (Werbespot, dargestellte Person, erwarteter Duft)

Für die positiven Merkmale des Werbespots ergibt sich für die drei Gruppen, die jeweils eine SF beurteilten, folgende Zusammensetzung der Merkmale, die in Abbildung 27 dargestellt ist. Die Graphen zeigen, dass sich SF II mit ihren Merkmalen der Stimme und Sprechweise (vgl. Kap. 6.1 und 6.2) in den Items edel, professionell und animierend von SF I und SF III deutlich absetzt und somit in der Wirkung der positiven Eigenschaften des Werbespots auf der Skala mit deutlich höheren Werten eingeschätzt wird. Die Differenz zu SF I beträgt dabei durchschnittlich 0.5 Punkte und zu SF III durchschnittlich 0.8 Punkte auf der Skala. Diese Tendenz deutet sich, wenngleich wesentlich schwächer ausgeprägt, auch bei dem Merkmal exklusiv an, wo SF II ebenfalls positiver bewertet wird als SF I und SF III.

#### Positive Merkmale des Werbespots



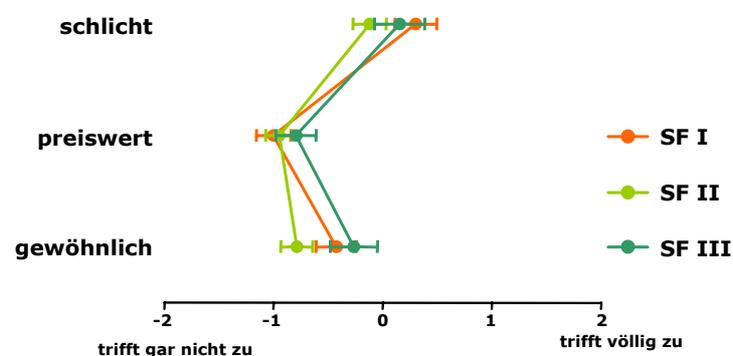
**Abb. 27: Graphische Darstellung der Mittelwerte  $\pm$  Standardfehler für die positiven Merkmale des Werbespots. Signifikante Unterschiede (Univariate Varianzanalyse, Post-Hoc-Test: Mehrfachvergleiche nach Tukey-HSD) zwischen SF I und SF II (\* für  $p \leq 0.5$ , \*\* für  $p \leq 0.1$ ) sowie SF II und SF III (# für  $p \leq 0.5$ , ## für  $p \leq 0.1$ ).**

Die statistische Auswertung ergibt, dass sich SF II mit Ausnahme des Merkmals exklusiv von SF I in den Merkmalen edel, professionell und animierend

signifikant ( $p \leq 0.5$ ), und von SF III sogar hoch signifikant ( $p \leq 0.1$ ) unterscheidet. Damit kann gezeigt werden, dass SF II eindeutig präferiert, und dass damit der gesamte Werbespot in dieser Gruppe, im Vergleich zu den anderen beiden, in seinen positiven Merkmalen als signifikant edeler, als signifikant professioneller und als signifikant animierender wahrgenommen wird.

Diese Tendenz lässt sich für die negativen Merkmale des Werbespots, die in Abbildung 28 dargestellt sind, nicht so eindeutig festhalten.

### Negative Merkmale des Werbespots



**Abb. 28: Graphische Darstellung der Mittelwerte  $\pm$  Standardfehler für die negativen Merkmale des Werbespots. Es ergeben sich keine signifikanten Unterschiede zwischen den SF (Univariate Varianzanalyse, Post-Hoc-Test: Mehrfachvergleiche nach Tukey-HSD).**

Abbildung 28 zeigt einen wesentlich homogenen Befund für die negativen Items des Werbespots zwischen den einzelnen SF. Besonders eng liegen die Urteile in Bezug auf das Merkmal preiswert beieinander (MW für SF I: -1.00, MW für SF II: -0.95 und MW für SF III: -0.80). Etwas differenzierter gestaltet sich das Bild für die Merkmale schlicht und gewöhnlich. Beide Merkmale werden zwar für SF II als am wenigsten zutreffend eingeschätzt, dennoch konnte statistisch keine Signifikanz in den Gruppenunterschieden nachgewiesen werden.

Als interessanter Befund sollen zudem die Ergebnisse angeführt werden, die sich aus den Gruppenunterschieden ergeben, wenn alle Merkmale des Werbespots (positive und negative) gemeinsam verglichen werden. Aus Tabelle 19 kann man entnehmen, dass sich für die Wirkung des gesamten Werbespots (Menge aller positiver und aller negativer Items) hochsignifikante Unterschiede zwischen den einzelnen Gruppen ergeben. Die Daten zeigen, dass sich dabei

SF II hochsignifikant von SF I ( $p \leq 0.01$ ) und ebenfalls hochsignifikant von SF III ( $p \leq 0.01$ ) unterscheidet und somit als signifikant wirkungsvoller in Bezug auf die Beurteilung des Werbespots herausgehoben werden kann. Aus den Daten ergibt sich damit auch, dass sich SF II als besonders wirkungsvoll von den anderen beiden SF absetzt und einheitlich präferiert wird.

Abhängige Variable: **Wirkung Werbespot gesamt** (positive und negative Merkmale)

(I) Gruppe	(J) Gruppe	Mittlere Differenz (I-J)	STAW	Signifikanz
SF I	SF II	- 0.40**	± 0.13	0.008**
	SF III	0.15	± 0.16	0.621
SF II	SF I	0.40**	± 0.13	0.008**
	SF III	0.54**	± 0.15	0.001**
SF III	SF I	- 0.15	± 0.16	0.621
	SF II	- 0.54**	± 0.15	0.001**

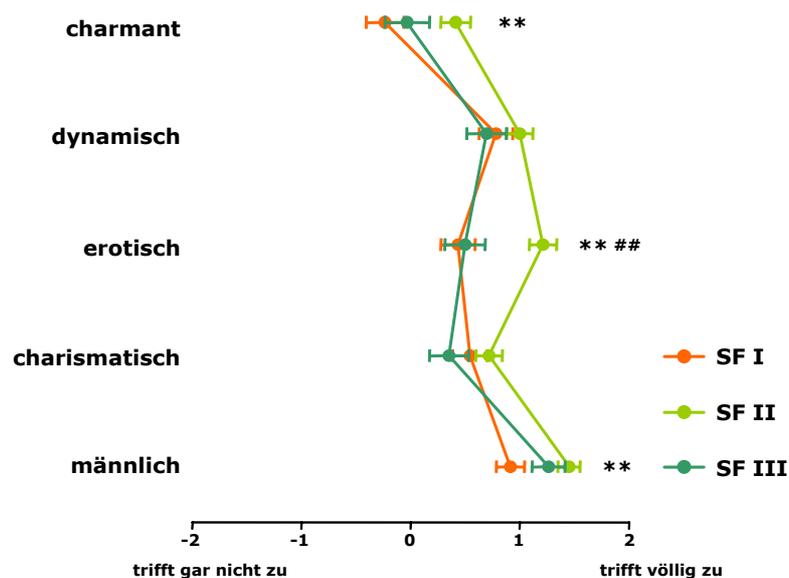
**Tab. 19: Tabellarische Darstellung der Mittleren Differenz, der Standardabweichung (STAW) und der daraus resultierenden Signifikanzen zwischen den einzelnen Gruppen für die Wirkung aller Items des gesamten Werbespots im Vergleich (Mehrfachvergleiche nach Tukey-HSD).**

In einem zweiten Block sollten die Probanden die Wirkung der im Werbevideo dargestellten Person einschätzen und bekamen dazu die Merkmale charmant, dynamisch, erotisch, weich, aufdringlich, charismatisch, männlich, abstoßend vorgelegt, die sie ebenfalls auf einer 5-er Skala von -2 (trifft gar nicht zu) bis +2 (trifft völlig zu) beurteilen sollten. Die Faktorenanalyse ergab auch bezüglich der Person eine Zuordnung der Items zu einem positiven Faktor (charmant, dynamisch, erotisch, charismatisch, männlich) und zu einem negativen Faktor (weich, aufdringlich, abstoßend).

Für die positiven Merkmale der Person ergibt sich die Darstellung, die in Abbildung 29 zusammengefasst ist. Die Daten zeigen, dass sich auch bei den positiven Merkmalen der Person SF II deutlich von SF I und SF III absetzt und vor allem bei den Merkmalen dynamisch (MW 1.00), erotisch (MW 1.21) und männlich (MW 1.45) hohe Werte, die im oberen Drittel der Skala angesiedelt sind, erreicht. Diese Tendenz lässt sich, etwas weniger stark ausgeprägt, allerdings auch für die Merkmale charmant und charismatisch beobachten, bei denen SF II ebenfalls durchschnittlich höhere Werte zu verzeichnen hat, als

die anderen beiden SF. Um diese Tendenzen statistisch zu stützen, wurden die Ergebnisse der einzelnen Gruppen durch einen Post-Hoc-Test (Tukey-HSD) miteinander verglichen. Als Resultat dieses Tests lässt sich festhalten, dass sich für die Merkmale charmant, erotisch und männlich jeweils hochsignifikante Unterschiede ( $p \leq 0.1$ ) zwischen SF I und SF II ergeben.

### Positive Merkmale der Person



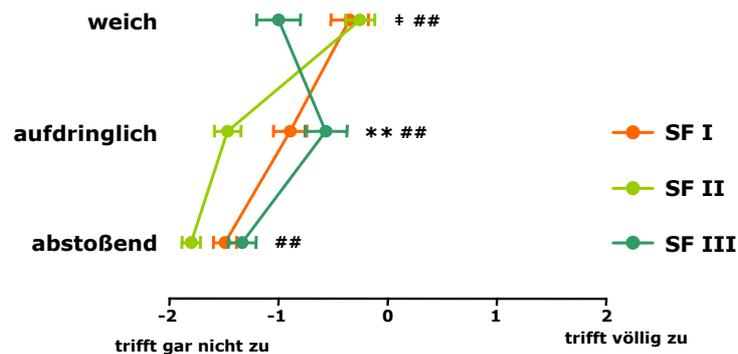
**Abb. 29: Graphische Darstellung der Mittelwerte  $\pm$  Standardfehler für die positiven Merkmale der Person. Signifikante Unterschiede (Univariate Varianzanalyse, Post-Hoc-Test: Mehrfachvergleiche nach Tukey-HSD) zwischen SF I und SF II (\* für  $p \leq 0.5$ , \*\* für  $p \leq 0.1$ ) sowie SF II und SF III (# für  $p \leq 0.5$ , ## für  $p \leq 0.1$ ).**

In diesen drei Items bewirkt SF II, dass die dargestellte Person des Werbespots als signifikant charmanter, signifikant erotischer und signifikant männlicher wahrgenommen wird, als dies in SF I der Fall ist. Darüber hinaus ergibt sich in Bezug auf das Merkmal erotisch auch ein hochsignifikanter Unterschied ( $p \leq 0.1$ ) zwischen SF II und SF III. Vor diesem Hintergrund lässt sich konstatieren, dass SF II von den Probanden einheitlich positiver bewertet, und somit als besonders wirkungsvolle Sprechfassung präferiert wird.

In Abbildung 30 ist nun abgebildet, wie die dargestellte Person des Werbespots bezüglich der negativen Merkmale zwischen den einzelnen SF beurteilt wird. Die Ergebnisse, die in Abbildung 30 veranschaulicht sind, zeigen zunächst ein etwas inhomogen wirkendes Bild zwischen den einzelnen SF. In An-

lehnung an die sehr hoch bewerteten positiven Eigenschaften der Person in SF II werden die negativen Merkmale aufdringlich (MW -1.47) und abstoßend (MW -1.8) entsprechend nah an der Grenze zu „trifft überhaupt nicht zu“ bewertet und reihen sich somit in ein konsistent wirkendes Bild ein.

### Negative Merkmale der Person



**Abb. 30: Graphische Darstellung der Mittelwerte  $\pm$  Standardfehler für die negativen Merkmale der Person. Signifikante Unterschiede (Univariate Varianzanalyse, Post-Hoc-Test: Mehrfachvergleiche nach Tukey-HSD) zwischen SF I und SF II (\* für  $p \leq 0.5$ , \*\* für  $p \leq 0.1$ ), SF II und SF III (# für  $p \leq 0.5$ , ## für  $p \leq 0.1$ ) sowie SF I und SF III (# für  $p \leq 0.5$ , ## für  $p \leq 0.1$ ).**

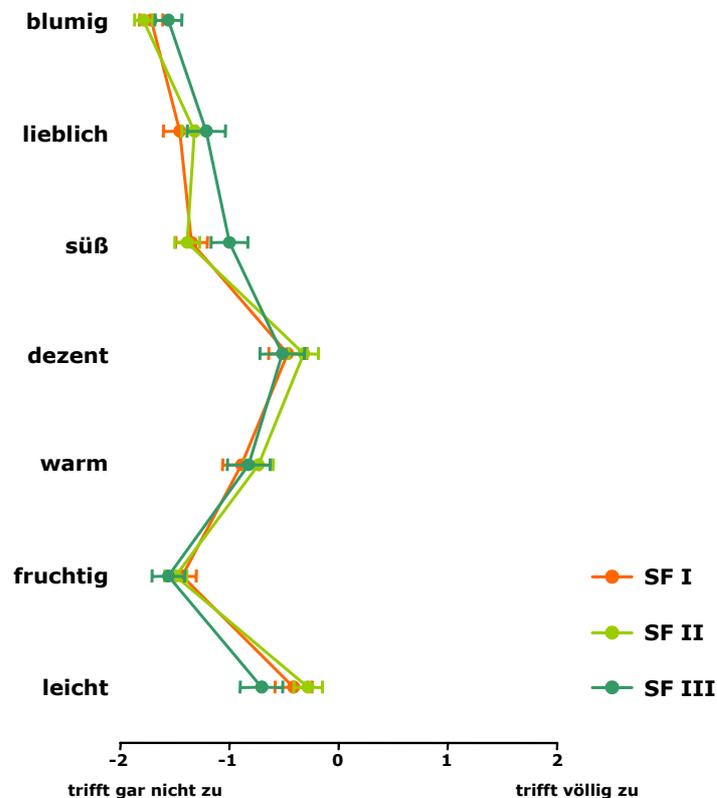
Die anderen beiden SF setzen sich in diesen Punkten deutlich von SF II ab und werden als entsprechend aufdringlicher bzw. abstoßender bewertet, wobei SF III jeweils die Spitzenposition einnimmt. Diese Tendenz konnte auch statistisch nachgewiesen werden. So unterscheiden sich SF II und SF III hochsignifikant ( $p \leq 0.1$ ) in den Merkmalen aufdringlich und abstoßend voneinander, was belegt dass SF III als besonders aufdringlich und abstoßend und SF II als besonders wenig aufdringlich und abstoßend von den Probanden wahrgenommen wird. In Bezug auf das Item aufdringlich ergibt sich zudem ein hochsignifikanter Unterschied ( $p \leq 0.1$ ) zwischen SF I und SF II, was ebenfalls unterstützt, dass SF II signifikant als weniger aufdringlich wahrgenommen, und dadurch als besonders wirkungsvolle SF präferiert wird.

Das Merkmal weich fällt aus den oben beschriebenen Zuordnungen tendenziell heraus. Hier setzt sich SF III als deutlich weniger weich bewertet von SF I und von SF II ab, die dicht beieinander liegen und, was auch die statistisch gewonnenen Daten belegen. Wie Abbildung 30 veranschaulicht, ergibt sich zwischen SF III und SF II ein hochsignifikanter Unterschied ( $p \leq 0.1$ ) und zwischen SF

III und SF I ein signifikanter Unterschied ( $p \leq 0.5$ ), was zur Folge hat, dass SF III als besonders wenig weich wirkend von den Probanden im Vergleich zu SF I und zu SF II eingeschätzt wird. Diese Einschätzung scheint in enger Anlehnung an die auditiven Eindrücke der Probanden gekoppelt, getroffen worden zu sein, was im folgenden Kapitel genauer diskutiert werden soll. Es gibt jedoch auch im akustischen Signal von SF III Indizien (vgl. Kap. 6.1 und 6.2), die von professionellen Hörern bestätigt wurden, und die eindeutig dafür sprechen, dass SF III als wenig weich im Vergleich zu den anderen beiden SF wahrgenommen wird.

In einem dritten Block sollten die Probanden eine große Anzahl an Attributen, die einen Duft charakterisieren und beschreiben können auf einer 5-er Skala von -2 (trifft gar nicht zu) bis +2 (trifft völlig zu) anhand des vorgeführten Werbevideos einschätzen. Dafür wurden folgende Eigenschaften vorgeschlagen: blumig, moschusartig, lieblich, markant, süß, tabakartig, dezent, warm,

#### Faktor: lieblich-fruchtig



**Abb. 31:** Graphische Darstellung der Mittelwerte  $\pm$  Standardfehler für die Dufteigenschaften des Faktors lieblich-fruchtig. Es ergeben sich keine signifikanten Unterschiede zwischen den SF (Univariate Varianzanalyse, Post-Hoc-Test: Mehrfachvergleiche nach Tukey-HSD).

herb, fruchtig, maskulin, leicht, elegant, frisch, holzig, sportlich, würzig, spritzig, klassisch, kühl, vitalisierend, und schwer.

Durch eine Faktorenanalyse ergaben sich für die Kategorie des Duftes die drei Faktoren (a) lieblich-fruchtig, (b) frisch-vitalisierend-sportlich und (c) würzig-holzig-herb, die nun in ihren Ergebnissen beschrieben werden sollen. Die Namensgebung der einzelnen Faktoren erfolgte jeweils durch die Attribute, die das Zentrum und damit den Kern eines jeden Faktors bildeten.

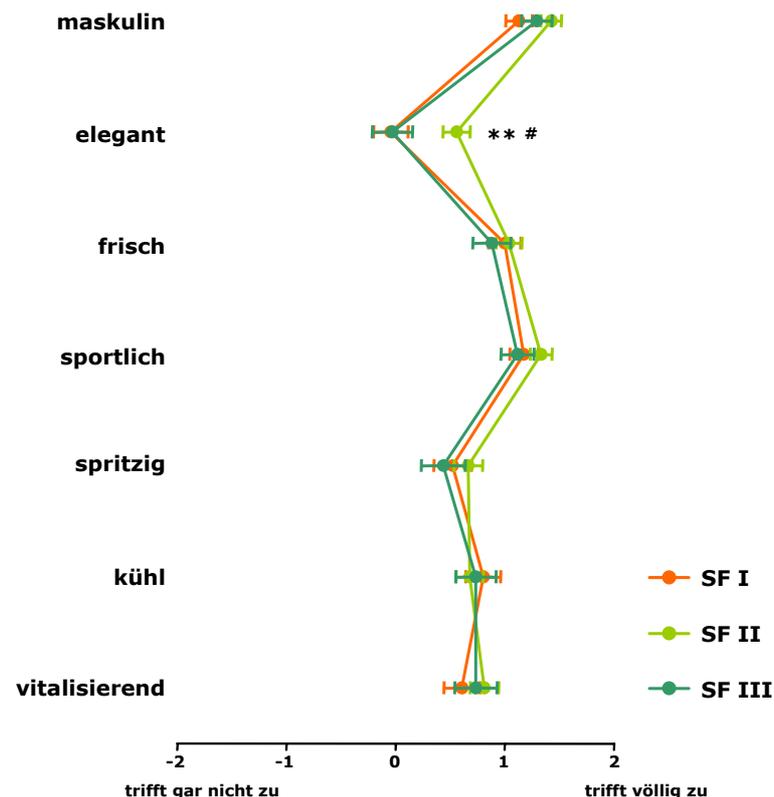
In Abbildung 31 sind die Werte für die Eigenschaften blumig, lieblich, süß, dezent, warm, fruchtig und leicht abgetragen und profilartig dargestellt, die zugleich den ersten Faktor ‚lieblich-fruchtig‘ charakterisieren. Die Profile der einzelnen SF weisen dabei ein durchweg homogenes Bild in der Beurteilung der einzelnen Items zwischen den verschiedenen Gruppen auf. Dabei liegen die MW für alle Eigenschaften durchweg im negativen Skalenbereich, so dass festgehalten werden kann, dass (a) durch statistische Testverfahren keine signifikanten Gruppenunterschiede nachgewiesen werden können und (b), dass unabhängig von der Sprechweise dieser Faktor als eher nicht zutreffend abgelehnt wird.

Demgegenüber stehen die Merkmale, die unter dem Faktor ‚frisch-vitalisierend-sportlich‘ gruppiert sind. Diesem Faktor lassen sich die Eigenschaften maskulin, elegant, frisch, sportlich, spritzig, kühl und vitalisierend zuordnen. Die grafische Darstellung in Abbildung 32 spiegelt ebenfalls einen sehr homogenen Befund zwischen SF I, SF II und SF III für nahezu alle Items wider. Im Gegenteil zu Faktor ‚lieblich-fruchtig‘ liegen für den Faktor ‚frisch-vitalisierend-sportlich‘ die Mittelwerte aller Merkmale allerdings über die drei SF hinweg durchgängig im positiven Skalenbereich. Dabei bilden die Merkmale maskulin, frisch und sportlich mit durchschnittlichen MW von 1.28, 0.97 und 1.26 den oberen Bereich der gegebenen Antworten zur Einschätzung des erwarteten Duftes überhaupt ab. Es kann von daher davon ausgegangen werden, dass der Werbespot, unabhängig von den einzelnen SF, einen besonders maskulinen, frischen und sportlichen Duft suggeriert, der von den Probanden sehr einheitlich assoziiert wird.

Aber auch die Merkmale spritzig, kühl und vitalisierend werden von den Rezipienten deutlich positiver bewertet, als alle zuvor beschriebenen Eigenschaften des Faktors ‚lieblich-fruchtig‘. Die einzige Ausnahme unter allen Charakteristika zur Beurteilung des Duftes bildet das Merkmal elegant. Hier konnte statis-

tisch nachgewiesen werden, dass sich SF II hochsignifikant ( $p \leq 0.1$ ) von SF I und signifikant ( $p \leq 0.5$ ) von SF III unterscheidet. Da es sich hierbei um ein einzelnes Phänomen handelt, wird dieser Erscheinung keine besondere Bedeutung beigemessen.

### Faktor: frisch-vitalisierend-sportlich

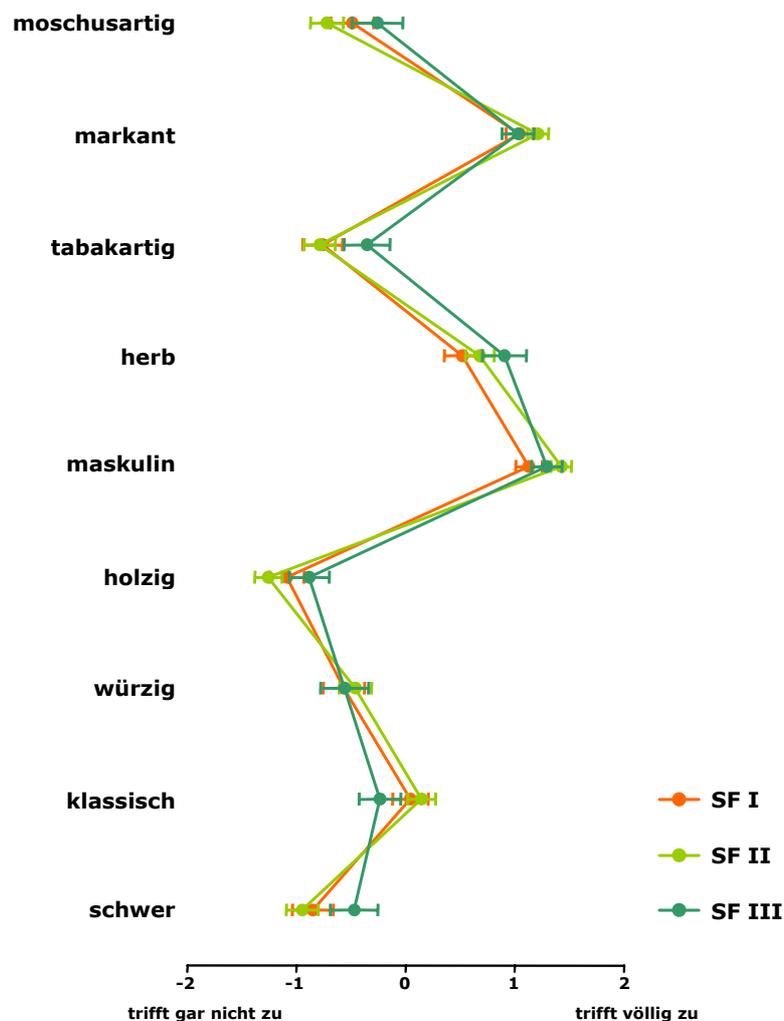


**Abb. 32: Graphische Darstellung der Mittelwerte  $\pm$  Standardfehler für die Dufteigenschaften des Faktors frisch-vitalisierend-sportlich. Signifikante Unterschiede (Univariate Varianzanalyse, Post-Hoc-Test: Mehrfachvergleiche nach Tukey-HSD) zwischen SF I und SF II (\* für  $p \leq 0.5$ , \*\* für  $p \leq 0.1$ ) sowie SF II und SF III (# für  $p \leq 0.5$ , ## für  $p \leq 0.1$ ).**

Neben den beiden bereits beschriebenen Faktoren, ergab sich durch die Faktorenanalyse ein weiteres Merkmalsbündel in Bezug auf den Duft, welches als Faktor ‚würzig-holzige-herb‘ in seinen Ergebnissen zusammenfassend in Abbildung 33 dargestellt ist. Dazu zählen die Eigenschaften moschusartig, markant, tabakartig, herb, maskulin, holzig, würzig, klassisch und schwer. Auch für die Items dieses Faktors werden sehr einheitliche Einschätzungen zwischen den einzelnen Gruppen getroffen, so dass die Graphen der drei SF einen sehr ähnlichen Kurvenverlauf aufweisen.

Differenzen zwischen dem Faktor ‚würzig-holzig-herb‘ und den anderen beiden Faktoren ergeben sich allerdings im Profil, welches sich aus den Verbindungslinien zwischen den separaten Merkmalen ergibt. So zeichnet sich der Faktor ‚würzig-holzig-herb‘ durch die Ausnutzung eines größeren Skalenbereiches aus. Die Eigenschaften markant, herb und maskulin werden beispielsweise durchgängig positiv bewertet, und liegen damit im oberen Drittel des Skalenbereiches, wohingegen Eigenschaften wie moschusartig, tabakartig, holzig und schwer als negativ im Sinne von eher nicht zutreffend beurteilt werden, und somit im unteren Drittel der Skala abgetragen sind.

### Faktor: würzig-holzig-herb



**Abb. 33: Graphische Darstellung der Mittelwerte  $\pm$  Standardfehler für die Dufteigenschaften des Faktors würzig-holzig-herb. Es ergeben sich keine signifikanten Unterschiede zwischen den SF (Univariate Varianzanalyse, Post-Hoc-Test: Mehrfachvergleiche nach Tukey-HSD).**

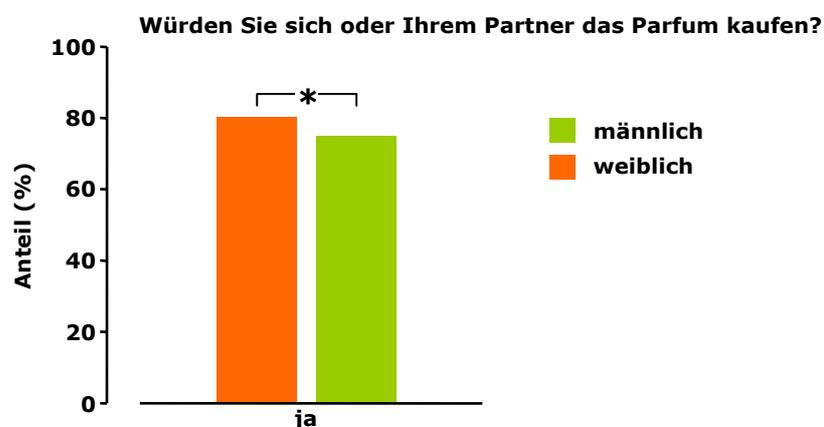
Für kein Merkmal des Faktors ‚würzig-holzig-herb‘ konnte durch statistische Tests ein signifikanter Gruppenunterschied zwischen den drei SF nachgewiesen werden, was wiederum ein Indiz für die sehr homogene Perzeption des Werbevideos bezüglich seiner Assoziation des erwarteten Duftes, unabhängig jedweder prosodischen Stimuli, darstellt.

### **Persuasion (Kaufeinstellung)**

Aufschlussreich in Bezug auf die Einschätzung der Wirksamkeit des Werbevideos gestalten sich auch die Ergebnisse, die die Einstellung der Probanden zum Testen bzw. Kaufen des beworbenen Parfums widerspiegeln. So wurden alle männlichen Probanden gebeten, die Frage, ob sie das Parfum testen würden mit ja oder mit nein zu beantworten, die Teilnehmerinnen der Untersuchung wurden gebeten, die Frage, ob sie ihrem Partner dieses Parfum kaufen würden, mit ja oder mit nein zu beantworten.

Dabei konnte ein signifikanter Zusammenhang zwischen der Geschlechterverteilung einerseits und dem Kaufverhalten andererseits nachgewiesen werden. Es kann gezeigt werden, dass prosodische Stimuli sowohl den perzeptiven Eindruck bezüglich der Wirkung einer Werbung, als auch die persuasiven Effekte (Kaufeinstellung, Einstellung zum Produkt) modifizieren können, was für die Werbeindustrie von vordergründigem Interesse sein dürfte.

Der Chi-Quadrat-Test nach Pearson ergibt, dass über die gesamte Stichprobe

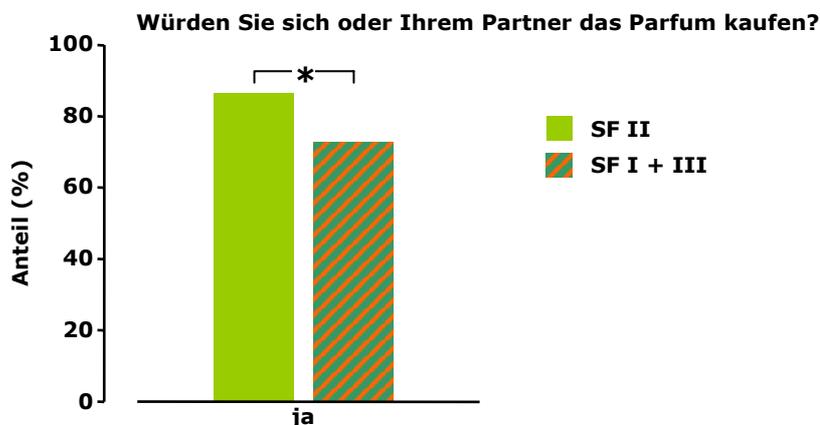


**Abb. 34: Graphische Darstellung des Zusammenhangs zwischen dem Geschlecht und dem Kaufverhalten (Chi-Quadrat-Test nach Pearson).**

hinweg, Frauen dem Kaufen des beworbenen Parfums gegenüber generell positiver eingestellt sind als Männer, was in Abbildung 34 verdeutlicht ist.

Des Weiteren wurde die Frage, ob die Probanden sich oder ihrem Partner das Parfum kaufen würden in Zusammenhang mit den einzelnen Sprechfassungen untersucht. Ziel der Frage war es, herauszufinden, ob sich die Kaufeinstellung oder das Interesse an dem Parfum durch die prosodische Merkmale der favorisierten Sprechfassung (SF II) ebenfalls positiv beeinflussen lassen, gegenüber den anderen beiden Sprechfassungen (SF I und SF III).

Durch statistische Testverfahren (Chi-Quadrat nach Pearson) kann gezeigt werden, dass es einen Zusammenhang zwischen der Einstellung und dem Kaufinteresse einerseits (Persuasion) und den prosodischen Merkmalen andererseits gibt. Wie auch für die oben beschriebenen Merkmale des Werbespots und der Person wird Sprechfassung II entgegen Sprechfassung I und III deutlich präferiert und zieht eine signifikant positivere Kaufeinstellung dem beworbenen Parfum gegenüber nach sich (Abb. 35).



**Abb. 35: Graphische Darstellung des Zusammenhangs zwischen dem Kaufverhalten und den drei Sprechfassungen (Chi-Quadrat-Test nach Pearson).**

Die beschriebenen Ergebnisse belegen, welche Bedeutung selbst kleine prosodische Variationen eines zudem kurzen Sprechanteils auf die Perzeption einer Werbung ausüben. Der Einfluss reicht dabei von Änderungen in der Wahrnehmung des Werbespots und der Person (Perzeption) bis zur veränderten Kaufeinstellungen (Persuasion) des Parfums. In Kapitel 6.4 werden diese Ergebnisse aller durchgeführten Analysen noch einmal zusammengefasst.

## 6.4 ZUSAMMENFASSUNG DER ERGEBNISSE

In einem Experiment, welches der Frage nach der Perzeption und Persuasion prosodischer Merkmale im Kontext einer Parfumwerbung nachging, konnte aufgezeigt werden, dass bereits kleine stimmliche und sprecherische Veränderungen innerhalb weniger Sekunden eine hochsignifikante Änderung in der Wahrnehmung eines Werbespots nach sich ziehen. Darüber hinaus konnte auch nachgewiesen werden, dass sich das Kaufverhalten und damit die Einstellung zum beworbenen Parfum durch prosodische Stimuli signifikant positiv verändern lassen.

Die Daten zeigen zudem, dass eine der untersuchten Sprechfassungen von allen Probanden einheitlich präferiert wird und somit auch im Sinne der Sprechwirkungsforschung generelle Akzeptanz in Bezug auf die kommunikative Adäquatheit für sich beanspruchen kann (Stock & Suttner 1991:69). Diese Sprechfassung (vgl. SF II) ist durch distinktive Eigenschaften charakterisiert, die sich sowohl apparativ als auch auditiv bestimmen lassen und die in Tabelle 20 aufgeführt sind.

- 
- **melodische Akzente** mit dezenten Melodiebewegungen **um große Terz** oder **Quarte** herum (keine übertriebenen Intervalle, aber auch keine monotone Sprechweise)
  - **Temporalität**, die sich durch eher **langsame Sprechweise** mit **angemessenen Pausen** auszeichnet
  - **dynamische Akzente**, die aus ihrem Umfeld heraustreten, aber nicht herausfallen (keine forcierten Intensitätsgipfel)
  - **sonore Basstimme**, die sich mit einem Durchschnittswert um **E** bzw. **F** unter dem für die Basstimmlage angegebenen Durchschnitt der Modalstimme bewegt
  - **keine Registerwechsel** zwischen Kopf- und Brustregister
  - **Knarren** nur **als Funktionsmittel**, anstelle des festen Stimmeinsatzes [ʔ], nicht als Wirkungsmittel
  - **Hauch** nur **als Funktionsmittel**, z. B. bei Aspiration von Plosiven, nicht als Wirkungsmittel
  - **hohe Stimmhaftigkeit** über die gesamte Äußerung (Modalstimme)
  - **weiches** und **warmes Timbre**, bei weder besonders heller, noch dunkler Klangfarbe der Stimme
- 

**Tab. 20: Überblick über distinktive Eigenschaften der Stimme und Sprechweise der präferierten Sprechfassung (SF II).**

In der Wirkungsuntersuchung konnte nun nachgewiesen werden, dass sich die aufgeführten Eigenschaften der Stimme und Sprechweise (Tab. 20) positiv auf die Perzeption und Persuasion des Werbespots auswirken, wobei der gesamte Werbespot in Sprechfassung II als hochsignifikant wirkungsvoller ( $p \leq 0.01$ ) im Vergleich zu den Sprechfassungen I und III eingeschätzt wurde und, wobei sich besonders die Eigenschaften edel, professionell und animierend als distinkte Wirkungsdeterminanten herauskristallisieren lassen.

Diese Tendenz lässt sich auch auf die Perzeption der dargestellten Person im Werbespot übertragen, die in SF II als hochsignifikant ( $p \leq 0.01$ ) charmanter, erotischer und männlicher im Vergleich zu Sprechfassung I und als ebenfalls hochsignifikant ( $p \leq 0.01$ ) erotischer im Vergleich zu Sprechfassung III eingeschätzt wird. Diese Einschätzung wird wiederum gestützt durch Daten, die sich aus der Beurteilung der negativen Items bezüglich der Person ergeben. Hier konnte gezeigt werden, dass SF II als hochsignifikant ( $p \leq 0.01$ ) weniger aufdringlich im Vergleich zu SF I und SF III und hochsignifikant ( $p \leq 0.01$ ) weniger abstoßend im Vergleich zu SF III eingeschätzt wurde.

Interessanterweise zeigen die Daten für die Beurteilung der einzelnen Attribute des erwarteten Duftes über die drei ermittelten Faktoren (a) lieblich-fruchtig, (b) frisch-vitalisierend-sportlich und (c) würzig-holzigherb hinweg keine signifikanten Gruppenunterschiede. Im Ergebnis kann gezeigt werden, dass die befragten Probanden, völlig unabhängig von den präsentierten Sprechfassungen, sehr einheitliche Bewertungen der einzelnen Duft-Items generieren. Darüber hinaus kristallisiert sich eine klare Bevorzugung aller Merkmale heraus, die in Faktor ‚frisch-vitalisierend-sportlich‘ gebündelt sind.

Die Ergebnisse zeigen zusammenfassend einen signifikanten Zusammenhang zwischen prosodischen Merkmalen einerseits und deren Perzeption (veränderte Wahrnehmung des Werbespots und der Person), bzw. deren Persuasion (veränderte KaufEinstellung gegenüber des Parfums) andererseits.

Im sich anschließenden Kapitel sollen die Daten der eigenen Untersuchung nun vor dem Hintergrund der aktuellen Forschungsliteratur kritisch diskutiert und eingeordnet werden.

## 7 DISKUSSION

In der vorliegenden Untersuchung sollte die Hypothese überprüft werden, dass Veränderungen von stimmlichen Parametern und sprecherischen Realisierungen eines Werbeslogans die Perzeption und die Persuasion eines Werbespots nachweislich beeinflussen. In einem Experiment wurden dazu drei verschiedene Sprechfassungen eines Werbeslogans durch messphonetische und auditive Analysemethoden und durch die Befragung von Probanden untersucht. Die Ergebnisse belegen eindeutig (vgl. Kap. 6), 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. Es lässt sich aus den Daten somit ein enger Zusammenhang zwischen prosodischen Merkmalen einerseits und deren Perzeption bzw. Persuasion andererseits belegen, womit die aufgestellte Hypothese der Arbeit verifiziert ist. Darüber hinaus verändert sich aber auch die Wahrnehmung einzelner Kategorien, z.B. die Wahrnehmung der dargestellten Person, wohingegen andere Kategorien, z.B. der erwartete Duft des Parfums von prosodischen Stimuli völlig unberührt bleiben. Über das Zusammenwirken prosodischer und olfaktorischer Stimuli gibt es bisher keine Studien, die eine Synästhesie der beiden Sinnesbereiche belegen, von daher ergäben sich hier mögliche Anknüpfungspunkte zu neuen Untersuchungen. Es lässt sich aber anhand der gewonnenen Ergebnisse die Vermutung anstellen, dass prosodische Marker einen direkten Einfluss auf konkret darstellbare Dimensionen (z.B. den Werbespot an sich, oder die dargestellte Person) ausüben, dass aber abstrakte Dimensionen (z.B. der Duft, der nur assoziiert, aber nicht bildlich wahrgenommen werden kann) von prosodischen Stimuli unerreicht bleiben.

In der Prosodieforschung spielen interkulturelle Unterschiede eine wesentliche Rolle und müssen bei der Beurteilung von Studien berücksichtigt werden. Nach Reinke (1998:191) sind Unterschiede in der Produktion und Perzeption emotionaler Sprechweisen auf bestimmte modellhafte Hörmuster zurückzuführen, was die Autorin unter dem Stichwort interkulturelle Emotionsrezeption zusammenfasst. Da einige der im Verlauf der Diskussion angeführten Studien aus dem angelsächsischen Raum stammen, kann davon ausgegangen werden, dass Unterschiede in der Bewertung einzelner stimmlicher Parameter oder Parameter der Sprechweise auch auf kulturelle Differenzen bezüglich der Hörge-

wohnheiten zurückzuführen sind und von daher nicht unkritisch auf die eigenen Daten übertragen werden können.

In Bezug auf das Encodieren von prosodischen Stimuli in Abhängigkeit vom Alter konnten Raithel & Hielscher-Fastabend (2004) in einer Untersuchung, die der Frage nach der Perzeption von affektiver und linguistischer Prosodie im Vergleich nachging, entgegen der Annahme von Breitenstein et al. (1996), aufzeigen, dass das Identifizieren und Zuordnen von affektiver Prosodie durch jüngere Probanden signifikant besser erfolgt. Da bei der eigenen Untersuchung 93.5% der Probanden unter 29 Jahre alt waren, kann davon ausgegangen werden, dass in der durchgeführten Studie affektive prosodische Phänomene zuverlässig erkannt wurden.

Mit dem Decodieren von prosodischen Parametern und deren Zuordnung zu verschiedenen Emotionen haben sich diverse Studien (Banse & Scherer 1996, Scherer et al. 1991, Bergmann et al. 1988, Ladd et al. 1985, Frick 1985, van Bezooijen 1984) beschäftigt. Dabei zeigen die Ergebnisse von Hörerurteilen in allen Studien, dass einzelne Emotionen allein auf der Basis von stimmlichen Merkmalen zuverlässig erkannt und zugeordnet werden können. Die Angaben über die prozentuale Genauigkeit für die exakte Zuordnung liegt dabei für die einzelnen Emotionen unterschiedlich hoch (vgl. Kap. 3.3.3). Da das Diskriminieren einzelner Emotionen nicht Gegenstand der eigenen Untersuchung ist, wird dieser Aspekt nicht weiter vertieft. Stattdessen sollen Ergebnisse, die einen Zusammenhang zwischen einzelnen prosodischen Parametern und konkreten affektiven Wirkungen aufzeigen, diskutiert werden. Banse & Scherer (1996) konnten zeigen, dass stimmliche Parameter nicht nur einen Index für die Aktivität (aktiv vs. passiv) von Emotionsdimensionen darstellen, sondern auch Aspekte der Valenz (angenehm vs. unangenehm) und der Qualität (positiv vs. negativ) widerspiegeln. Der Versuch, Emotionen durch die drei Dimensionen Aktivität, Valenz und Qualität abzubilden, stellt eine Taxonomie dar, die es neben der Darstellung gradueller Unterschiede zwischen einzelnen Emotionen auch erlaubt, Unterschiede zwischen Emotionskategorien aufzuzeigen. Die drei Emotionsdimensionen werden in Anlehnung an Wundt (1905, 1911) erstmals bei Schlosberg<sup>15</sup> (1954) genannt. Vor dem Hintergrund dieser Emotionsdimensionen verweisen Pereira (2000) und Schröder et al. (2001) in ihren Ar-

---

<sup>15</sup> In der Literatur finden sich für die Bezeichnung der drei Emotionsdimensionen differierende Termini, wie z.B. arousal, valence, control oder activation, evaluation, power (vgl. Kap. 2.2.1).

beiten darauf, dass sich Emotionen mit ähnlichen Erregungsgraden, wie z.B. Angst und Ärger, gleiche akustische Parameter teilen, und deswegen bei entsprechenden Tests durch Hörurteile oft nicht gut voneinander differenziert werden können. Untersuchungen von Ladd et al. (1985) und Schröder et al. (2001) belegen anhand einzelner Parameter ( $F_0$ -Verlauf, mittlere Grundfrequenz, Intensität, Pausen, Stimmqualität), dass die häufigsten und stärksten Korrelationen dieser Parameter mit der Aktivitätsdimension einhergehen. Demnach gehen aktive Emotionen vor allem mit einer erhöhten Grundfrequenz, kürzeren Pausen, stärkeren und schnelleren  $F_0$ -Anstiegen bzw. Abfällen und einer ansteigenden Intensität einher. Für die Valenz- und Qualitätsdimension sind Korrelationen zwischen prosodischen Parametern und der Perzeption von Emotionen ebenfalls, aber weniger stark ausgeprägt, existent. Ladd et al. (1985) führen beispielsweise an, dass raue Stimmen bezüglich ihrer Valenz eher mit negativen Dimensionen assoziiert werden.

Die drei Dimensionen Valenz, Qualität und Aktivität (vgl. Scherer 2003 und Kap. 3.1) lassen sich auch auf die Ergebnisse der eigenen Untersuchung übertragen. In einem Versuch (vgl. Kap. 5 und Kap. 6) sollten die Probanden in Form eines Fragebogens das Video einer Parfumwerbung bezüglich des Werbespots, der dargestellten Person und des assoziierten Duftes beurteilen. Drei separate Gruppen bekamen dabei jeweils unterschiedliche Videos, die sich ausschließlich in ihrer Sprechweise unterschieden, wobei die Probanden in keiner Weise vorab für prosodische Phänomene sensibilisiert wurden.

In der Zuordnung der Kategorien des Fragebogens zu den Emotionsdimensionen ergeben sich Parallelen, die sich gemäß der Methode des semantischen Differenzials (Osgood 1952 und Osgood et al. 1957) durch binär angeordnete Koordinaten abbilden lassen, was in Tabelle 21 verdeutlicht ist. Daraus geht hervor, dass die Attribute, die zur Analyse des Werbespots erhoben wurden, mit der Qualitätsdimension einhergehen. Dabei ergeben sich zwei Faktoren, die statistisch durch eine Faktorenanalyse (vgl. Kap. 5.2) bestätigt wurden und die sich den der Qualität entsprechenden Dimensionen positiv und negativ zuordnen lassen. Darüber hinaus zeigt sich, dass prosodische Phänomene vor allem auf die positive Dimension einen entscheidenden Einfluss ausüben. Welche einzelnen Stimmerkmale für die positive Perzeption dabei besonders wirkungsvoll sind, zeigen die Ergebnisse der apparativen und auditiven Analy-

se, die an späterer Stelle ausführlicher diskutiert werden (vgl. Kap. 6.1 und Kap. 6.2).

FRAGEBOGENKATEGORIE	EMOTIONSDIMENSION		
<b>Werbespot</b>	<b>Qualität</b>		
		<b>positiv</b> (exklusiv, edel, professionell, animierend)	<b>negativ</b> (schlicht, preiswert, gewöhnlich)
<b>Person</b>	<b>Aktivität und Qualität</b>		
	<b>aktiv positiv</b> (dynamisch)	<b>passiv positiv</b> (charmant, erotisch, männlich)	<b>passiv negativ</b> (weich, aufdringlich, abstoßend)
<b>Duft</b>	<b>Valenz</b>		
	<b>angenehm</b> (maskulin, elegant, frisch, sportlich, spritzig, kühl, vitalisierend)	<b>unangenehm</b> (blumig, lieblich, süß, dezent, warm, fruchtig, leicht)	

**Tab. 21: Zusammenhang zwischen den Kategorien des Fragebogens und den Emotionsdimensionen. Die grün unterlegten Spalten kennzeichnen die Bereiche, in denen prosodische Merkmale signifikante Unterschiede zwischen den verschiedenen Sprechweisen nach sich ziehen und damit eine signifikante Veränderung in der Perzeption und Wirkung des Werbespots markieren.**

Um die Attribute zur Charakterisierung der dargestellten Person hinreichend systematisieren zu können, muss ein Merkmalsbündel aus den Emotionsdimensionen Aktivität und Qualität gebildet werden. Die Unabhängigkeit der Skalen zur Abbildung der Emotionsdimensionen hat Izard (1999) untersucht und dabei existierende Korrelationen nachgewiesen. Durch die Verbindung zweier Skalen ergeben sich drei Kategorien für die Bewertung der Person. Wie man Tabelle 21 entnehmen kann, sind alle Eigenschaften, die von ihrer Aktivität als passiv und von ihrer Qualität aber sowohl als positiv als auch negativ in das Konzept der Emotionsdimensionen eingeordnet werden können, sehr empfänglich für prosodische Stimuli. Die Bevorzugung der eher passiv anmutenden emotionalen Färbungen lassen sich aus dem Kontext des Werbevideos und durch die Bilder, die Schnitte, die Blenden und die gleichförmig verlaufende Musik erklären. Dass in diesem Zusammenhang besonders eine Sprechweise

präferiert wird, die diese Zurückgenommenheit synästhetisch aufgreift, erweist sich als gelungene Synthese.

Ordnet man den Attributen zur Charakterisierung des Duftes Emotionsdimensionen zu, so lassen sich die Duftassoziationen bezüglich ihrer Valenz in angenehm und unangenehm wirkende Gerüche untergliedern. Diese Einteilung erfolgte statistisch über die gesamte Stichprobe hinweg sehr homogen, so dass wie man Tabelle 21 entnehmen kann prosodische Merkmale keinen Einfluss auf die Veränderung der Perzeption bezüglich des Duftes ausüben.

In allen Fällen, in denen prosodische Merkmale zu signifikanten Veränderungen der Perzeption führten, wurden Merkmale der Stimme und der Sprechweise einer Sprechfassung (SF II) präferiert, so dass im Ergebnis festgehalten werden kann, dass die prosodischen Merkmale dieser Sprechfassung wirkungsvoller sind als jene der anderen beiden untersuchten Sprechfassungen. Scherer et al. (1973) sprechen in diesem Zusammenhang von einer ‚voice of confidence‘. Es handelt sich in dieser Sprechfassung (II) um Merkmale der Akzentuierung aus einer melodischen Komponente (dezente Melodiebewegungen um eine große Terz bzw. Quarte herum), einer dynamischen Komponente (keine forcierten Intensitätsgipfel) und einer temporalen Komponente (eher langsame Sprechweise und angemessene Pausen), der Sprechstimmlage (sopran Bassstimme um E bzw. F herum), der Register (Modalstimme ohne Registerwechsel), der Stimmqualität (Knarren und Hauch nur als Funktions-, nicht als Wirkungsmittel) und des Timbres (weiche und warme Klangfarben). Wie sich diese Merkmale wiederum in den Zusammenhang mit den Ergebnissen anderer Studien bringen lassen, soll im Folgenden diskutiert werden.

In einer Untersuchung von Cauldwell (2000), die der Frage nach der Wirkung prosodischer Stimuli in kontextgebundenen und isolierten Situationen im Vergleich nachging, konnte als Ergebnis aufgezeigt werden, dass Emotionalität, wenn die prosodischen Merkmale isoliert präsentiert werden, eine stärkere Wirkung nach sich zieht als in vergleichbarer Präsentation im Kontext. Dieses Ergebnis besitzt bezüglich der eigenen Fragestellung unter methodischen Gesichtspunkten Relevanz, als in der eigenen Untersuchung prosodische Phänomene eines Werbespots im Kontext mit visuellen Reizen (Bilder) und akustischen Reizen (Musik) präsentiert wurden, was von der Überlegung, eine möglichst authentischen Untersuchungssituation zu erzeugen, geleitet war. Bezogen auf die Ergebnisse der eigenen Untersuchung scheint es aber umso ein-

drucksvoller, dass zusätzliche Stimuli wie Musik und Bild die Wirkung prosodischer Impulse nicht nivelliert haben, wie es die Ergebnisse Cauldwells (2000) zunächst nahe legen. Darüber hinaus wäre es in einer gesonderten Studie interessant zu untersuchen, wie sich die Wirkung prosodischer Merkmale in der Werbung bei Präsentation in isolierter Situation unter Gesichtspunkten der Perzeption und Persuasion veränderte.

Verschiedene Arbeiten haben sich mit der Perzeption von temporalen Phänomenen der Sprechgeschwindigkeit und der Pausensetzung beschäftigt und dabei enge Beziehungen zueinander herauskristallisiert. Die Interessenschwerpunkte der Untersuchungen waren entweder geleitet von der Frage danach, ob Veränderungen in der Perzeption untersuchter Äußerungen an die kognitive Aktivität der Rezipienten gebunden sind (Goldman-Eisler 1968, Huggins 1972, Lehiste 1975, Meinhold 1995), oder von der Frage danach, ob Veränderungen in der Wahrnehmung von der Sprecherpersönlichkeit abhängig sind (Smith et al. 1975, Miller et al. 1976, Apple et al. 1979, Steward & Ryan 1982, Woodall & Burgoon 1983, Sharf & Lehman 1984, Brown et al. 1985, Burgoon et al. 1990, Peterson et al. 1995).

In Zusammenhang mit der Bewertung einer Persönlichkeit kommen die oben genannten Arbeiten einheitlich zu dem Ergebnis, dass ein erhöhtes Sprechtempo auch die Glaubwürdigkeit (Miller et al. 1976, Burgoon et al. 1990) eines Sprechers erhöht, ihn darüber hinaus wohlgesonnener (Steward & Ryan 1982), kompetenter (Brown et al. 1985, Burgoon et al. 1990) und erfolgreicher (Sharf & Lehman 1984) erscheinen lässt und seine persuasiven Effekte (Woodall & Burgoon 1983) vergrößert.

Langsameres Sprechen hingegen wird als weniger ehrlich, weniger gewandt und passiver wahrgenommen (Apple et al. 1979). Diese Ergebnisse, die größtenteils in Studien aus dem angelsächsischen Raum erhoben wurden und als Zielgruppe erfolgreiche Verkäufer und Interviewer fokussierten, lassen sich nicht bedenkenlos auf die eigenen Ergebnisse übertragen, die die Temporalität im Zusammenhang mit einer Parfumwerbung im deutschen Sprachraum betrachtet. Darüber hinaus werden in den meisten Arbeiten längere Gesprächssequenzen untersucht, wohingegen in der eigenen Untersuchung ein Werbeslogan analysiert wird, der über die drei Sprechfassungen (SF) hinweg eine Länge von 3.35s (SF I), 4.03s (SF II) und 4.19s (SF III) umfasst.

Interessanterweise konnte sich in der eigenen Analyse gerade eine eher langsam und damit auch passiv anmutende Sprechweise als besonders wirkungsvoll durchsetzen, was sich wiederum gut in den Kontext des gesamten Werbespots, der eher minimalistisch und zurückgenommen anmutet, einordnen lässt. Smith et al. (1975) konnten in einer Untersuchung nachweisen, dass es zwar einerseits einen linearen Zusammenhang zwischen der Kompetenz eines Sprechers und der Sprechgeschwindigkeit gibt (schnelleres Sprechen wirkt kompetenter), dass aber zwischen der Absicht des Erzeugens von Wohlwollen (benevolence) eines Sprechers und dessen Sprechgeschwindigkeit ein U-förmiger Verlauf existiert, wobei höchstes Wohlwollen mit gemäßigter Sprechgeschwindigkeit einhergeht.

Es stellt sich die Frage, wie solche minimalen Unterschiede in der Temporalität, wie sie für die drei Sprechfassungen beschrieben sind, überhaupt eine Änderung in der Wahrnehmung nach sich ziehen. Eine mögliche Erklärung für dieses Phänomen liegt im Zusammenspiel verschiedener prosodischer Parameter. So weist Sprechfassung I, die in ihrer zeitlichen Struktur kürzer ist als die anderen beiden Sprechfassungen, zudem eine große Beweglichkeit in der Melodisierung auf. Einen positiven Zusammenhang dieser Interaktion beschreibt Trouvain (2004), indem er feststellt, dass durch größere Tonhöhendifferenzen auch der perzeptive Eindruck einer größeren Sprechgeschwindigkeit entsteht. Dieser Zusammenhang konnte in einer entsprechenden Analyse durch das auditive Urteil von geschulten Hörern bestätigt werden.

Sprechfassungen II und III liegen mit durchschnittlichen Werten von 3.88 Silben pro Sekunde bzw. 3.27 Silben pro Sekunde weit unter dem Durchschnitt der Erzählgeschwindigkeit, den Goldman-Eisler (1968) mit 4.4-5.9 Silben pro Sekunden angibt. Darüber hinaus wird Sprechfassung II ( $t_{\text{ges}}$  4.03s) von geschulten Hörern in einer auditiven Analyse als langsamer bezüglich der Sprechgeschwindigkeit beurteilt als Sprechfassung III ( $t_{\text{ges}}$  4.19s). Dieses kontroverse Urteil zwischen apparativer und auditiver Analyse lässt sich vor dem Hintergrund der Ergebnisse Goldman-Eislers (1968) erklären, die längere Pausen als Pendant für eine langsamere Sprechgeschwindigkeit angibt. Im Werbeslogan „Caractère | **PAUSE** | Macht am Mann | PAUSE | das Markante am Mann.“ variiert die erste Pause, die gleichsam den Markennamen und den Slogan voneinander trennt, zwischen SF II mit 0.66s und SF III mit 0.24s stark, wobei die zweite Pause nahezu identische Realisationszeiten zwischen

den Sprechfassungen (0.36s vs. 0.37s) aufweist. Daraus resultiert eine durchschnittliche Pausenzeit, die mit 14.6% für Sprechfassung III deutlich unter der von Sprechfassung II mit 25.3% liegt, so dass die Feststellung Goldman-Eislers (1968), wonach längere Pausen mit dem perzeptiven Eindruck einer langsameren Sprechgeschwindigkeit korrelieren, durch die Ergebnisse der eigenen Analyse bestätigt werden. Überdies liegt die Vermutung nahe, dass perzeptive Urteile bezüglich der Temporalität bereits durch sehr kleine Sprechereinheiten etabliert und nachhaltig manifestiert werden. Diese Hypothese könnte Gegenstand einer separaten Untersuchung sein.

Ein weiterer Parameter, der in zahlreichen Arbeiten unter verschiedenen Fragestellungen in Bezug auf die Perzeption betrachtet wird, ist die Grundfrequenz (mittlerer  $F_0$ -Wert und  $F_0$ -Verlauf). Dabei stellt sich die Frage danach, inwieweit prosodische Parameter bezüglich ihrer Perzeption überhaupt separierbar sind. Ladd et al. (1985) finden in einer Untersuchung heraus, dass prosodische Parameter weitgehend unabhängig voneinander den Höreindruck beeinflussen. Die Arbeiten, die sich mit Phänomenen der Melodisierung beschäftigen, nähern sich der Problematik unter drei Gesichtspunkten: (a) im Zusammenhang mit der Bedeutung für Sprachsyntheseverfahren (Schröder et al. 2001, Schröder 2004, Burkhardt & Sendlmeier 1999), (b) im Zusammenhang mit dem Decodieren von Emotionen (Williams & Stevens 1972, Frick 1986, Bergmann et al. 1988, Murray & Arnott 1993, Kienast 2002, Paeschke 2003, Kranich 2003) und (c) im Zusammenhang mit der Identifikation einer Sprecherbefindlichkeit (Scherer et al. 1973, Apple et al. 1979, Sharf & Lehman 1984, Burgoon et al. 1990, Peterson et al. 1995), wobei letztgenannter Punkt für die eigene Fragestellung die meisten Anknüpfungspunkte bietet. Die Ergebnisse bilden jedoch ein eher inhomogenes Bild ab, was mit großer Wahrscheinlichkeit auch auf interkulturelle Differenzen zurückzuführen ist. Bezüglich der Variabilität der Grundfrequenz sind größere  $F_0$ -Bewegungen Indiz hoher Aktivierung (Schröder et al. 2001) und spiegeln demnach Ärger und Erregung (Bergmann et al. 1988) oder Kompetenz, Charakter und Geselligkeit, bzw. in Kombination mit anderen Parametern auch eine erhöhte Persuasivität und Glaubwürdigkeit (Burgoon et al. 1990) wider. Peterson et al. (1995) nennen stark abfallende Melodiekurven am Ende einer Äußerung als Ursache für eine autoritäre, glaubwürdige und dominante Ausstrahlung.

Zu konträren Ergebnissen bezüglich der mittleren Sprechstimmlage kommen Apple et al. (1979), die herausfinden, dass eine manipulierte Grundfrequenz, die um 20% erhöht ist, einen Sprecher als wenig glaubwürdig, wenig emphatisch, wenig potent, klein, dünn und nervös erscheinen lässt, wohingegen Sharf & Lehman (1984) eine relativ hohe Grundfrequenz als Indiz für erfolgreiche Interviewer anführen. Letztere haben für ihre Studie ausschließlich weibliche Probanden gewählt, so dass die abweichenden Ergebnisse möglicherweise auch darauf zurückgeführt werden können.

Mit den Ergebnissen der eigenen Wirkungsanalyse konnte nachgewiesen werden, dass unter Gesichtspunkten der Melodisierung eine tiefe sonore Bassstimme präferiert wird, die mit einer mittleren Sprechstimmlage um E bzw. F herum, ca. eine Terz unter den für den Durchschnitt angegebenen Werten der Bassstimme liegt (vgl. Kap. 4.1). Zudem konnte sich eine Sprechweise durchsetzen, die sich durch dezente Melodiebewegungen (um eine große Terz, bzw. Quarte herum) auszeichnet, und dabei weder übertrieben große Intervalle, noch eine monotone Kurvenführung aufweist. Da in den meisten Untersuchungen keine genauen Angaben über exakte Intervalle gemacht werden, lassen sich die Daten leider nur ungefähr zueinander in Beziehung setzen.

In einer frühen Arbeit von Scherer et al. (1973) können die Autoren aufzeigen, dass eine prosodisch zweifelhafte Sprechweise (große  $F_0$ -Bewegungen und erhöhte  $F_0$ ) wenig wirkungsvoll ist, und dass es etwas wie eine „voice of confidence“ gibt, die auf paralinguistisch stimmige Parameter zurückzuführen ist, was für die eigene Untersuchung durch die oben benannten Stimmmerkmale bestätigt werden kann.

Vom Forschungsdesign kommt eine Studie von Gélinas-Chebat & Chebat (1992) der eigenen Untersuchung am nächsten. Die Autoren gehen der Frage nach, inwieweit Veränderungen der Melodisierung und der Intensität in der Werbung, die Einstellung zu einem beworbenen Produkt verändern können. Dabei finden sie heraus, dass wenig Kontur in der  $F_0$ -Kurve eine stärkere Persuasion nach sich zieht, was die Ergebnisse der eigenen Untersuchung ebenfalls belegen. Darüber hinaus beschreiben sie anhand des Modells der Elaborationswahrscheinlichkeit (vgl. Kap. 2.2) einen größeren Einfluss des  $F_0$ -Verlaufs auf die Hörer, wenn die Rezipienten am semantischen Inhalt der Werbung nicht interessiert sind, von dem sie überzeugt werden sollen. Da in der eigenen Untersuchung darauf geachtet wurde, dass auf semantischer Ebene keine

persuasiven Elemente (Fakten oder Argumente) enthalten sind, kann davon ausgegangen werden, dass die Merkmale der Sprechstimmlage (tiefe, sonore Basstimme) und der Melodisierung (Melodiebewegungen um eine große Terz bzw. Quarte) einen entscheidenden Einfluss auf die Präferenz der beschriebenen Sprechfassung ausgeübt haben, die sich als besonders wirkungsvoll herauskristallisiert hat.

Einen weiteren Parameter, der in der Literatur als maßgebliche Wirkungsdeterminante in Zusammenhang mit persuasiven und perceptiven Eindrücken beschrieben wird, stellt die Intensität dar. Dabei korreliert eine steigende durchschnittliche Intensität bei männlichen Sprechern positiv mit der Aktivitätsdimension (Pereira 2000, Schröder et al. 2001). Studien, die Probanden aus dem angelsächsischen Sprachraum untersucht haben, kommen einheitlich zu dem Ergebnis, dass ein erhöhtes Intensitätslevel positiver bewertet wird (Sharf & Lehman 1984), und dass eine dynamischere Sprechweise, hervorgerufen u.a. durch eine größere Lautstärke, einen Sprecher überzeugender und glaubhafter erscheinen lässt (Pearce & Brommel 1972, Burgoon et al. 1990) als einen vergleichbaren Sprecher mit konventioneller Sprechweise. Bergmann et al. (1988) stellen diesen Befunden eine Studie gegenüber, in der eine hohe Intensität kennzeichnend für negative Emotionen und Einstellungen mit einem hohen Erregungsgrad (Aggressivität, Ärger, Vorwurf) ist, was die Daten der eigenen Untersuchung ebenfalls belegen. Hier wurde die Sprechfassung mit dem höchsten Intensitätsgrad (SF III) als signifikant aufdringlicher und abstoßender empfunden und damit in ihrer Wirkung abgelehnt, was sich wiederum wie für die anderen Parameter auch aus dem Kontext des gesamten Werbespots heraus (vgl. Kap. 5 und Kap. 6) erklären lässt.

Als abschließende Wirkungsparameter sollen noch die Stimmqualität und die Artikulationspräzision angeführt werden, die durch das Zusammenspiel der Muskelaktivität eng miteinander einhergehen (vgl. Kap. 4.1). Dabei wird eine unterspannte Artikulationspräzision eher von einer behauchten Stimmqualität und eine überspannte Artikulationspräzision von einer verknarnten oder gepressten Stimmqualität begleitet. In den meisten Arbeiten werden die Spannungszustände an der Glottis selbst oder oberhalb der Glottis in Zusammenhang mit dem Erregungsgrad eines Sprechers und dem Decodieren von Emotionen untersucht (Murray & Arnott 1993, Tolkmitt & Scherer 1986, Kienast & Sendlmeier 2000, Gobl & Chasaide 2000). Unter Gesichtspunkten der Stimm-

wirkung postulieren Ladd et al. (1985), dass raue Stimmen mit dem Erregungsgrad eines Sprechers interferieren und Befindlichkeiten wie Ärger und Belästigung charakterisieren.

In den drei Sprechfassungen, die in dieser Arbeit im Wirkungsprozess einer Parfumwerbung untersucht wurden, konnte klar aufgezeigt werden, dass Hauch, Knarren und Pressen als Wirkungsmittel der Stimme deutlich zugunsten einer sonoren Modalstimme, die sich durch eine durchgängig hohe Stimmhaftigkeit auszeichnet, abgelehnt werden.

Wie die angeführten Arbeiten belegen, handelt es sich im Wirkungsprozess prosodischer Merkmale um ein komplexes Bündel an Parametern, die weitestgehend zusammenwirken, wobei einzelne Merkmale aus dem Kontinuum heraustreten können und somit zu Veränderungen von Sprechweisen in den Dimensionen Aktivität, Qualität und Valenz führen. Unter den untersuchten Sprechfassungen dieser Arbeit wurde eine Sprechfassung (SF II) präferiert. Sie ist damit im Sinne der Sprechwirkungsforschung (Stock & Suttner 1991) auch besonders wirkungsvoll, und lässt sich mit ihren prosodischen Merkmalen anhand der Einteilung in die Emotionsdimensionen nach Scherer (vgl. Abb. 4) am deutlichsten in den oberen Bereich des Quadranten positiv und passiv einordnen. Mit der Präferenz dieser Attribute kann auch aufgezeigt werden, dass die von den Rezipienten bevorzugte Sprechweise die elegante Zurückgenommenheit des Werbespots synästhetisch aufgreift (Auswahl des Protagonisten, der Musik, der Schnitte und der Blenden) und sich dadurch exzellent in den gesamten Kontext der Parfumwerbung einordnet.

Eine positive Interferenz von Prosodie und Kontext auf der einen Seite und Perzeption auf der anderen Seite haben Schirmer et al. (2002, 2005, 2005a) neurophysiologisch untersucht und dabei herausgefunden, dass affektive Prosodie beim Decodieren von Inhalten zusätzliche Informationsquellen enthält, die allerdings geschlechtsspezifisch unterschiedlich genutzt werden. Im Zusammenhang mit der Sprachverarbeitung und mit der Wortverarbeitung nutzen Frauen prosodische Stimuli effektiver und schneller als Männer (Schirmer et al. 2002, 2005, 2005a). Eine Analyse der Daten des eigenen Korpus unter Gesichtspunkten der Varianz innerhalb der Geschlechterverteilung bietet sich als separate Fragestellung für eine weiterführende Untersuchung an. Es konnte über die gesamte Stichprobe hinweg gezeigt werden, dass die Gruppe der Frauen generell dem Kauf des beworbenen Parfums gegenüber positiver ein-

gestellt war, als die Gruppe der Männer. Inwieweit sich diese Einstellung innerhalb der Untergruppen jedoch durch prosodische Merkmale geschlechtsspezifisch ändert, konnte aufgrund einer zu inhomogenen Zusammensetzung der einzelnen Teilgruppen nicht genauer charakterisiert werden.

## **8 ZUSAMMENFASSUNG UND AUSBLICK**

In der vorliegenden Arbeit wurde die dichotome Betrachtung prosodischer Phänomene unter formellen und funktionellen Gesichtspunkten im Zusammenhang mit der Frage nach der Persuasion und Perzeption stimmlicher und sprecherischer Parameter in einer Parfumwerbung untersucht.

Ausgehend von aktuellen Forschungsarbeiten, die sich mit ausgewählten prosodischen Parametern und deren Funktion im Wirkungsprozess beschäftigen, wurde in Anlehnung an Scherers Adaption (1978, 2000a) des Linsenmodells von Brunswik (1956) ein experimentelles Design entwickelt, das phonetische und sprechwissenschaftliche Ansätze miteinander verbindet. In einem Experiment, in dem Rezipienten den Werbespot einer Parfumwerbung unter verschiedenen Aspekten beurteilen sollten, ohne dabei für prosodische Phänomene sensibilisiert zu werden, konnte durch eine groß angelegte Fragebogenuntersuchung deutlich aufgezeigt werden, dass es so etwas wie eine „voice of confidence“ (Scherer et al. 1973) gibt, die in ihrer Wirkung eine signifikante Differenzierung in der Betrachtung des Werbespots bezüglich seiner Perzeption (Wahrnehmung) und seiner Persuasion (Einstellung zum Produkt und zum Kauf des Produktes) nach sich zieht. Diese Fassung, die sich durch eine ruhige Sprechweise, durch eine tiefe und sonore Basstimme, durch eine hohe Stimmhaftigkeit ohne Hauch und Knarren und durch dezente Melodiebewegungen auszeichnet, steht mit dem gesamten Werbespot in der Auswahl der Bilder, der Schnitte, der Blenden, der Musik und der dargestellten Person in synästhetischem Einklang. Der Werbespot in seiner Gesamtheit zeichnet sich dabei durch eine schlichte Eleganz aus.

Die Ergebnisse sind in dieser Detailliertheit und Prägnanz bisher in keiner anderen vergleichbaren Studie vorgelegt worden.

Darüber hinaus haben sich durch die Diskussion der eigenen Daten vor dem Hintergrund anderer Studien Anregungen ergeben, die Gegenstand weiterführender Arbeiten sein könnten. So stellt sich beispielsweise die Frage, inwieweit die Perzeption von prosodischen Stimuli in der Werbung kontextgebunden ist und sich damit bei isolierter Präsentation oder in fremde Kontexte eingebettet, verändert. Interessant erscheint auch die Frage danach, wie sich die Perzeption verändert, wenn man gezielt einzelne Parameter synthetisch manipuliert, und ob es Parameter gibt, die einen stärkeren Einfluss auf die Perzeption und

Persuasion haben als andere. Nicht zuletzt interessiert auch die Frage danach, ab wann Veränderungen einzelner prosodischer Parameter überhaupt den perceptiven Eindruck beeinflussen und auf welches Minimum prosodische Impulse herabgesetzt werden können, um dennoch persuasive Effekte zu konstituieren. Möglicherweise bieten die vorliegenden Befunde darüber hinaus Ansatzpunkte zur Untersuchung der neuropsychologischen Verarbeitung von Prosodie in der Werbung, was z.B. mit Hilfe der funktionellen Magnetresonanztomographie interdisziplinär weiter analysiert werden könnte.

## 9 LITERATURVERZEICHNIS

- ABADJIEVA, E./I. R. MURRAY/J. L. ARNOTT** (1993): Applying Analysis of Human Emotional Speech to Enhance Synthetic Speech. In: Proceedings of the 3<sup>rd</sup> European Conference on Speech Communication and Technology (Eurospeech), Bd. 2, S. 909-912.
- ACKERMANN, H./I. HERTRICH/W. ZIEGLER** (1993): Prosodische Störungen bei neurologischen Erkrankungen – eine Literaturübersicht. In: Fortschritte der Neurologie, Psychiatrie und ihrer Grenzgebiete 61, Heft 7, S. 241-253.
- ACKERMANN, H./I. HERTRICH/W. GRODD/D. WILDGRUBER** (2004): „Das Hören von Gefühlen“: Funktionell-neuroanatomische Grundlagen der Verarbeitung affektiver Prosodie. In: Aktuelle Neurologie 31, S. 449-460.
- ADDINGTON, D. W.** (1971): The Effect of Vocal Variations on Ratings of Source Credibility. In: Speech Monographs 38, Heft 1, S. 242-247.
- ALTMANN, H.** (Hrsg.) (1988): Intonationsforschungen. Tübingen: Niemeyer.
- AMIR, N./S. RON/N. LOAR** (2000): Analysis of an emotional speech corpus in Hebrew based on objective criteria. In: Proceedings of the ISCA Workshop on Speech and Emotion.  
<http://www.qub.ac.uk/en/isca/proceedings>
- ANDERS, L. C.** (1997): Spektrale Analysen gestörter Stimmen. Habilitationsschrift. Jena.
- ANDERS, L. C.** (2000): Klassifizierungssysteme zur Stimmklangbewertung in der klinischen Praxis. In: Geißner, H. (Hrsg.) Stimmen hören. 2. Stuttgarter Stimmtage. St. Ingbert: Röhrig, S. 21-28.
- ANDERS, L. C.** (2001): Bewertung von Ergebnissen der Stimmfunktionstherapie – physiologische und auditive Aspekte. In: Bräunlich, M./B. Neuber/B. Rues (Hrsg.) Gesprochene Sprache – transdisziplinär. Frankfurt a. M. [u. a.]: Lang, S. 9-18.
- ANDERS, Y.** (2001a): Merkmale der Melodisierung und des Sprechausdrucks ausgewählter Dichtungsinterpretationen im Urteil von Hörern. Frankfurt a. M. [u.a.]: Lang.
- APPLE, W./L. A. STREETER/R. M. KRAUSS** (1979): Effects of Pitch and Speech Rate on Personal Attributions. In: Journal of Personality and Social Psychology 37, Heft 5, S. 715-727.
- ARDENER, E.** (Hrsg.) (1970): Social Anthropology and Language. Association of Social Anthropologists Monograph 10, London: Tavistock Press.
- BANNERT, R.** (1985): Toward a Model for German Prosody. In: Folia Linguistica 19, S. 321-341.

- BANSE, R./K. R. SCHERER** (1996): Acoustic Profiles in Vocal Emotion Expression. In: Journal of Personality and Social Psychology 70, Heft 3, S. 614-636.
- BARRY, W. J.** (Hrsg.) (2004): Phonus 7, Saarbrücken: Institut für Phonetik Universität des Saarlandes.
- BARRY, W. J.** (Hrsg.) (2004): Phonus 8, Saarbrücken: Institut für Phonetik Universität des Saarlandes.
- BASTIAN, H.-J.** (1986): Zur Bedeutung des limbischen Systems und der Theorie der Emotionen für die Stimmwirkungsforschung. In: Höhle, T./E.-M. Krech/G. Lerchner/D. Sommer (Hrsg.): Hallesche Studien zur Wirkung von Sprache und Literatur 12. Wissenschaftliche Beiträge der Universität Halle 1986/28 (F61), S. 42-48.
- BENKWITZ, A.** (2004): Kontrastive phonetische Untersuchungen zum Rhythmus. Britisches Englisch als Ausgangssprache – Deutsch als Zielsprache. Frankfurt a. M. [u. a.]: Lang.
- BERGER, L.** (Hrsg.) (1984): Sprechausdruck. Frankfurt a. M.: Scriptor.
- BERGMANN, G.** (1984): Vokale Kommunikation: Methoden, Ergebnisse und Probleme der Erforschung des Sprechausdrucks. In: Berger, L. (Hrsg.): Sprechausdruck. Frankfurt a. M.: Scriptor, S. 69-82.
- BERGMANN, G./T. GOLDBECK/K. R. SCHERER** (1988): Emotionale Eindruckswirkung von prosodischen Sprechmerkmalen. In: Zeitschrift für experimentelle und angewandte Psychologie 35, Heft 2, S. 167-200.
- BEZOOIJEN, R. VAN** (1984): The characteristics and recognizability of vocal expression of emotions. Dordrecht: Foris.
- BIEGE, A./I. BOSE** (1998) (Hrsg.): Theorie & Empirie in der Sprechwissenschaft. Hanau [u.a.]: Dausien.
- BOHNER, G.** (2003): Einstellungen. Kapitel 8. In: Stroebe, W./K. Jonas/M. Hewstone (Hrsg.): Sozialpsychologie. Eine Einführung. <sup>4</sup>Berlin [u.a.]:Springer, S. 265-315.
- BOLINGER, D.** (1986): Intonation and its parts. Stanford: Stanford University Press.
- BOLINGER, D.** (1989): Intonation and its uses. Stanford: Stanford University Press.
- BOROD, J.** (Hrsg.) (2000): The neuropsychology of Emotion. Oxford [u.a.]: Oxford Univ. Press.
- BOSE, I.** (2001): Methoden der Sprechausdrucksbeschreibung am Beispiel kindlicher Spielkommunikation. In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion 2, S. 262-303.

- BOSE, I.** (2003): dóch da sín ja' nur mûster//. Kindlicher Sprechausdruck im sozialen Rollenspiel. Frankfurt a. M. [u.a.]: Lang.
- BOWER, A. B.** (2001): Highly Attractive Models in Advertising and the Women Who Loathe Them: The Implications of Negative Affect for Spokesperson Effectiveness. In: *The Journal of Advertising* 30, Heft 3, S. 51-63.
- BRÄUNLICH, M./B. NEUBER/B. RUES** (2001) (Hrsg.) *Gesprochene Sprache – transdisziplinär*. Frankfurt a. M. [u. a.]: Lang.
- BREITENSTEIN, C./I. DAUM/H. ACKERMANN/R. LÜTGEHETMANN/E. MÜLLER** (1996): Erfassung der Emotionswahrnehmung bei zentralnervösen Läsionen und Erkrankungen: Psychometrische Gütekriterien der „Tübinger Affekt Batterie“. In: *Neurologische Rehabilitation* 2, S. 93-101.
- BROWN, R. W./A. H. BLACK/A. T. HOROVITZ** (1955): Phonetic symbolism in natural languages. In: *Journal of abnormal and social psychology* 50, S. 388-393.
- BROWN, B. L./H. GILES/J. N. THAKERAR** (1985): Speaker Evaluations as a Function of Speech Rate, Accent, and Context. In: *Language and Communication*, 5 (3), S. 207-220.
- BROWN, B. L./J. M. BRADSHAW** (1985): Towards a Social Psychology of Voice Variations. In: Giles, H./R. N. Clair (Hrsg.): *Recent Advances in Language, Communication, and social Psychology*. London: Erlbaum, S. 144-181.
- BRUNNER, A.** (2000): Suprasegmentalia aus phonetisch/phonologischer Sicht. In: Geißner, H. (Hrsg.) *Stimmen hören*. 2. Stuttgarter Stimmtage. St. Ingbert: Röhrig, S. 151-152.
- BRUNSWIK, E.** (1956): *Perception and the Representative Disign of Psychological Experiments*. Berkeley: University of California Press.
- BURGOON, J. K./T. BRINK/M. PFAU** (1990): Nonverbal Behaviors, Persuasion, and Credibility. In: *Human Communication Research* 17, Heft 1, S. 140-169.
- BURKHARDT, F./W. F. SENDLMEIER** (1999): Simulation der Emotion "Freude" mit Sprachsyntheseverfahren. In: *Forum Acusticum. Collected Papers of the 137<sup>th</sup> Meeting of the Acoustical Society of America and the 2<sup>nd</sup> Convebtion of the European Acoustics Association Integrating the 25<sup>th</sup> German Acoustics DAGA Conference*. Berlin.
- BUBMANN, H.** (Hrsg.) (1990): *Lexikon der Sprachwissenschaft*. <sup>2</sup>Stuttgart: Kröner.
- CANDLAND, D. K.** (Hrsg.) (1962): *Emotion: Bodily change*. Princeton: Van Nostrand.

- CAULDWELL, R. T.** (2000): Where Did The Anger Go? The Role Of Context In Interpreting Emotion In Speech. In: Proceedings of the ISCA Workshop on Speech and Emotion.  
<http://www.qub.ac.uk/en/isca/proceedings>
- CHAIKEN, S./A. LIBERMANN/A. H. EAGLY** (1989): Heuristic and systematic information processing within and beyond the persuasion context. In: Uleman, J. S./J. A. Bargh (Hrsg.): Unintended thought. New York: Guilford, S. 212-252.
- CHEBAT, J. C./D. VAILLANT/C. GÉLINAS-CHEBAT** (2000): Does background music enhance sales persons' persuasiveness? In: Perceptual and Motor Skills 91, S. 405-424.
- CHILDS, G. T.** (1994): African ideophones. In: Hinton, L./J. Nichols/J. J. Ohala (Hrsg.): Sound symbolism. Cambridge: Cambridge University Press, S. 178-204.
- CRYSTAL, D.** (1970): Prosodic and paralinguistic correlates of social categories. In: Ardener, E. (Hrsg.): Social Anthropology and Language. Association of Social Anthropologists Monograph 10, London: Tavistock Press, S. 185-206.
- CRYSTAL, D.** (1975): The English tone of voice. London: Edward Arnold.
- DARWIN, C. R.** (1901): Der Ausdruck der Gemüthsbewegungen bei dem Menschen und den Thieren: mit 21 Holzschnitten und 7 Tafeln. <sup>5</sup>Stuttgart: Schweizerbart.
- DAVITZ, J. R.** (1964): The Communication of Emotional Meaning. McGraw-Hill: New York.
- DEMPWOLFF, O.** (1919): Über phonetische Aichung. In: VOX: Mitteilungen aus dem Phonetischen Laboratorium Hamburg 94. Hamburg [u.a.].
- EAGLY, A. H./S. CHAIKEN/W. WOOD** (1981): An Attribution Analysis of Persuasion. In: Harvey, J. H./W. J. Ilkes/R. F. Kidd (Hrsg.): New Directions in Attribution Research. Hillsdale: Erlbaum, S. 37-62.
- ECKERT, H./J. LAVER** (1994): Menschen und ihre Stimmen. Aspekte der vokal- len Kommunikation. Weinheim: Beltz.
- EKMAN, P./W. v. FRIESEN** (1971): Constants Across Cultures In The Face And Emotion. In: Journal of Personality and Social Psychology 17, Heft 2, S. 124-129.
- EKMAN, P.** (1984): Expression and the Nature of Emotion. In: Scherer, K. R./P. Ekman (Hrsg.): Approaches to Emotion. Hillsdale [u.a.]: Erlbaum, S. 319-343.
- EKMAN, P.** (1992): An argument for basic emotions. In: Cognition and Emotion 6, Heft 3/4, S. 169-200.

- ESSEN, O. VON** (1952): Melodien deutscher Dichtung. In: Essen, O. von/C. Gutknecht/J.-P. Köster (Hrsg.): Sprechmelodie als Ausdrucksgestaltung. Hamburger Phonetische Beiträge. Hamburg: Buske, S. 1-24.
- ESSEN, O. VON/C. GUTKNECHT/J.-P. KÖSTER** (Hrsg.): Sprechmelodie als Ausdrucksgestaltung. Hamburger Phonetische Beiträge. Hamburg: Buske.
- ESSEN, O. VON** (1956): Grundzüge der hochdeutschen Satzintonation. Ratingen [u.a.]: Henn.
- ESSEN, O. VON** (1979): Allgemeine und angewandte Phonetik. <sup>5</sup>Berlin: Akademie-Verlag.
- ERTEL, S.** (1969): Psychophonetik: Untersuchungen über Lautsymbolik und Motivation. Göttingen: Hogrefe.
- ETZEL, S.** (1983): Untersuchungen zur Lautsymbolik. Inauguraldissertation der Johann Wolfgang Goethe-Universität, Frankfurt a. M.
- FÄHRMANN, RUDOLF** (1967): Die Deutung des Sprechausdrucks. Studien zur Einführung in die Praxis der charakterologischen Stimm- und Sprechanalyse. <sup>2</sup>Bonn: Bouvier.
- FELSER, G.** (2001): Werbe- und Konsumentenpsychologie. Eine Einführung. Stuttgart: Schäffer-Poeschel.
- FERNANDEZ, R./R. PICARD** (2000): Modelling drivers' speech under stress. In: Proceedings of the ISCA Workshop on Speech and Emotion. <http://www.qub.ac.uk/en/isca/proceedings>
- FINE, S. H.** (1992): Marketing the Public Sector: Promoting the Causes of Public and Nonprofit Agencies. New Brunswick: Transaction Publishers.
- FRICK, R. W.** (1985): Communicating Emotion: The Role of Prosodic Features. In: Psychological Bulletin 97, Heft 3, S. 412-429.
- FRICK, R. W.** (1986): The Prosodic Expression of Anger: Differentiating Threat and Frustration. In: Aggressive behavior 12, S. 121-128.
- FRIEDRICH, G./W.BIGENZAHN** (1995): Phoniatrie. Einführung in die medizinischen, psychologischen und linguistischen Grundlagen von Stimme und Sprache. Bern [u.a.]: Hans Huber.
- FRINDTE, WOLFGANG** (2001): Einführung in die Kommunikationspsychologie. Weinheim [u.a.]: Beltz.
- FRÖHLICH, W. D.** (2000): Wörterbuch Psychologie. <sup>23</sup>München: DTV.
- GEIBNER, H.** (1969): Rede in der Öffentlichkeit – Eine Einführung in die Rhetorik. Stuttgart [u.a.] Kohlhammer.
- GEIBNER, H.** (1981): Sprechwissenschaft. Theorie der mündlichen Kommunikation. Königstein: Scriptor.

- GEIBNER, H.** (2000) (Hrsg.) Stimmen hören. 2. Stuttgarter Stimmtage. St. Ingbert: Röhrig.
- GEIBNER, U.** (1985): Lehrerreaktionen und sprecherischer Ausdruck. Zur Relevanz von suprasegmentalen Merkmalen in Unterrichtsprozessen. Gießen: Schmitz.
- GÉLINAS-CHEBAT, C./J. C. CHEBAT** (1992): Effects of Two Characteristics on the Attitudes Toward Advertising Messages. In: The Journal of Social Psychology 132, Heft 4, S. 447-459.
- GÉLINAS-CHEBAT, C./J. C. CHEBAT/A. VANINSKY** (1996): Voice and Advertising Effects on Intonations and Intensity of Voice on Source Credibility, Attitudes toward the Advertized Service and the Internet to Buy. In: Perceptual and Motor Skills 83, S. 243-262.
- GÉLINAS-CHEBAT, C./J. C. CHEBAT** (1999): Impact of Voice on Source Credibility in Advertising: A self-Monitoring Approach. In: North American Journal of Psychology 1, Heft 2, S. 323-342.
- GIBBON, D./H. RICHTER** (1984) (Hrsg.) Intonation, accent and rhythm: Studies in discourse phonology. Berlin [u.a.]: de Gruyter.
- GIBBON, D.** (1988): Intonation and Discourse. In: Petöfi, J. S. (Hrsg.): Text and Discourse Constitution. Empirical Aspects, Theoretical Approaches. Berlin [u.a.]: de Gruyter, S. 3-25.
- GILES, H./R. N. CLAIR** (1985) (Hrsg.): Recent Advances in Language, Communication, and social Psychology. London: Erlbaum.
- Gless, D.** (1988) (Hrsg.): Stimmbildung und Liedgestaltung. Berlin: Volk und Wissen.
- GLÜCK, H.** (2000) (Hrsg.): Metzler Lexikon Sprache. Stuttgart [u.a.]: Metzler.
- GOBL, C./A. N. CHASAIDE** (2000): Testing Affective Correlates Of Voice Quality Through Analysis And Resynthesis. In: Proceedings of the ISCA Workshop on Speech and Emotion.  
<http://www.qub.ac.uk/en/isca/proceedings>
- GOLDMAN-EISLER, F.** (1968): Psycholinguistics. Experiments in Spontaneous Speech. London [u.a.]: Academic Press.
- GORN, G.** (1982): The effects of music in advertising on chose behaviour: A classical conditioning approach. In: Journal of Marketing Research 46, S. 94-101.
- GÖTTERT, K.-H.** (1998): Geschichte der Stimme. München: Fink.
- GREASLEY, P./C. SHERRARD/M. WATERMAN** (2000): Emotion in language and speech: Methodological issues in naturalistic approaches. In: Language and Speech 43, Heft 4, S. 355-375.
- GRIMM, H./J. ENGELKAMP** (1981): Sprachpsychologie. Berlin: Schmidt.

- GUNDERMANN, H.** (1994): Phänomen Stimme. München [u.a.]: Reinhardt.
- GUNDERMANN, H.** (Hrsg.) (1998): Die Ausdruckswelt der Stimme. 1. Stuttgarter Stimmtage. Heidelberg: Hüthig.
- GÜTTLER, P. O.** (1996): Sozialpsychologie: soziale Einstellungen, Vorurteile, Einstellungsänderungen. <sup>2</sup>München [u.a.] Oldenbourg.
- HALL, J. A.** (1980): Voice Tone and Persuasion. In: Journal of Personality and Social Psychology 38, Heft 6, S. 924-934.
- HALLIDAY, M. A. K.** (1967): Intonation and grammar in British English. The Hague [u.a.]: Mouton.
- HALLIDAY, M. A. K.** (1995): An Introduction to Functional Grammar. <sup>2</sup>London [u.a.]: Arnold.
- HART, J. 'T/R. COLLIER/A. COHEN** (1990): A perceptual study of intonation. Cambridge: Cambridge University Press.
- HARVEY, J. H./W. J. ILKES/R. F. KIDD** (1981) (Hrsg.): New Directions in Attribution Research. Hillsdale: Erlbaum.
- HEIDOLPH, K. E./W. FLÄMIG/W. MOTSCH** und ein Autorenkollektiv (1980): Grundzüge einer deutschen Grammatik. Berlin: Akademie Verlag.
- HEIKE, G.** (1969): Suprasegmentale Analyse. Marburg: Elwert.
- HELMS, S.** (1981): Musik in der Werbung. Wiesbaden: Breitkopf & Härtel.
- HEWSTON, M./W.STROEBE/G. M. STEPHENSON** (Hrsg.) (1996): Introduction to Social Psychology. <sup>2</sup>Oxford [u.a.]: Blackwell.
- HINTON, J./J. NICHOLS/J. J. OHALA** (Hrsg.) (1994): Sound Symbolism. Cambridge: Cambridge University Press.
- HIRSCHFELD, U.** (1995): Kleines phonetisches Wörterbuch. In: Fremdsprache Deutsch. Zeitschrift für die Praxis des Deutschunterrichts 1, Heft 12, S. 56.
- HIRSCHFELD, U.** (1998): Probleme des Sprachvergleichs im Bereich der Intonation. In: Biege, A./I. Bose (Hrsg.): Theorie & Empirie in der Sprachwissenschaft. Hanau [u.a.]: Dausien, S. 75-84.
- HÖHLE, T./E.-M. KRECH/G. LERCHNER/D. SOMMER** (Hrsg.) (1986): Hallesche Studien zur Wirkung von Sprache und Literatur 12. Wissenschaftliche Beiträge der Universität Halle 1986/28 (F61).
- HOLLMACH, U.** (1996): Soziophonetische Grundlagen zur Neukodifizierung des Aussprachewörterbuches. In: Krech, E.-M./Stock, E. (Hrsg.): Beiträge zur deutschen Standardaussprache. Hanau [u.a.]: Dausien, S. 60-67.
- HOVLAND, C. I./L. J. IRVING/H. H. KELLEY** (Hrsg.) (1966): Communication and Persuasion. Psychological Studies of Opinion Change. <sup>10</sup>New Haven: Yale University Press.

- HUGGINS, A. W. F.** (1972): On the Perception of Temporal Phenomena in Speech. In: The Journal of the Acoustical Society of America 51, Bd. 4 (2), S. 1279-1290.
- IZARD, C. E.** (1999): Die Emotionen des Menschen. Eine Einführung in die Grundlage der Emotionspsychologie. <sup>4</sup>Weinheim: Beltz.
- JORDAN, W.** (1870): Durch's Ohr. Mirosfiche-Ausgabe 1990-1994. München [u.a.]: Saur.
- KEHREIN, R.** (2002): Prosodie und Emotionen. Tübingen: Niemeyer.
- KENT, R. D./M. J. BALL** (Hrsg.) (2000): Voice quality measurement. San Diego: Singular.
- KIENAST, M./W. F. SENDLMEIER** (2000): Acoustical Analysis and Temporal changes in Emotional Speech. In: Proceedings of the ISCA Workshop on Speech and Emotion.  
<http://www.qub.ac.uk/en/isca/proceedings>
- KIENAST, M.** (2002): Phonetische Veränderungen in emotionaler Sprechweise. Aachen: Shaker.
- KLASMEYER, G./W. F. SENDLMEIER** (2000): Voice and emotional states. In: Kent, R. D./M. J. Ball (Hrsg.): Voice quality measurement. San Diego: Singular, S. 339-357.
- KLEINGINNA, P. R./A. M. KLEINGINNA** (1981): A categorized list of emotion definitions, with suggestion for a consensual definition. In: Motivation and Emotion 5, S. 345-379.
- KLOß, I.** (2000): Werbung: Lehr-, Studien- und Nachschlagewerk. <sup>2</sup>München [u.a.]: Oldenbourg.
- KNAPE, J.** (2003): Persuasion. In: Ueding, G.: Historisches Wörterbuch der Rhetorik, Bd. 6. Niemeyer: Tübingen, S. 874-907.
- KNOWLES, G.** (1984): Variable Strategies in Intonation. In: Gibbon, D./H. Richter (Hrsg.) Intonation, accent and rhythm: Studies in discourse phonology. Berlin [u.a.]: de Gruyter, S. 226-242.
- KOHLER, K. J.** (1977): Einführung in die Phonetik des Deutschen. Berlin: Schmidt.
- KOHLER, K. J.** (Hrsg.) (1998): Arbeitsberichte 33, Kiel: Institut für digitale Sprachverarbeitung.
- KÖHLER, W.** (1929): Gestalt Psychologie. New York: Mentor Books.
- KRAMER, E.** (1963): Judgement of Personal Characteristics and Emotions from Nonverbal Properties of Speech. In: Psychological Bulletin 60, S. 408-420.

- KRANICH, W.** (2003): Phonetische Untersuchungen zur Prosodie emotionaler Sprechausdrucksweisen. Frankfurt a. M.: Lang.
- KRECH, E.-M./G. RICHTER/E. STOCK/J.SUTTNER** (Hrsg.) (1991): Sprechwirkung. Grundfragen, Methoden und Ergebnisse ihrer Erforschung. Berlin: Akademie-Verlag.
- KRECH, E.-M./STOCK, E.** (1996) (Hrsg.): Beiträge zur deutschen Standardausprache. Hanau [u.a.]: Dausien.
- KRECH, H.** (1960): Einführung in die deutsche Sprechwissenschaft/Sprecherziehung. Berlin: Deutscher Verlag der Wissenschaften.
- KROEBER-RIEL, W.** (1996): Bildkommunikation. Imagerystrategien für die Werbung. München: Vahlen.
- Kröger, B. J./C. Riek/G. Sachse** (1998) (Hrsg.): Festschrift für Georg Heike. Frankfurt a. M.: Hector.
- LADD, D. R./K. E. A. SILVERMAN/F. TOLKMITT/G. BERGMANN/K. R. SCHERER** (1985): Evidence for the independent function of intonation contour type, voice quality, and  $F_0$  range in signaling speaker affect. In: The Journal of the Acoustical Society of America 78, Bd. 4, S. 435-444.
- LADEFOGED, P.** (1971): Preliminaries to linguistic phonetics. Chicago: University of Chicago Press.
- LAVER, J.** (1980): The phonetic description of voice quality. London [u.a.]: Cambridge University Press.
- LEHISTE, I.** (1976): Influence of fundamental frequency pattern on the perception of duration. In: Journal of Phonetics 4, S. 113-117.
- LÖHKEN, S. C.** (1997): Deutsche Wortprosodie. Abschwächungs- und Tilgungsvorgänge. Studien zur deutschen Grammatik 56. Tübingen: Stauffenburg.
- LUCHSINGER, R./G. E. ARNOLD** (1970): Handbuch der Stimm und Sprachheilkunde. Die Stimme und ihre Störungen. Bd. 1. Wien [u.a.]: Springer.
- MEES, U.** (1991): Die Struktur der Emotionen. Göttingen [u.a.]: Hogrefe.
- MEID, V.** (1999): Sachwörterbuch zur deutschen Literatur. Stuttgart: Reclam.
- MEINHOLD, G.** (1995): Zeitparameter gesprochener Sprache. Forschungsbericht. Universität Jena. Lehrstuhl für Phonetik und Sprechwissenschaft (Manuskript).
- MEYER, W.-U./A. SCHÜTZWOHL/R. REISENZEIN** (1993): Einführung in die Emotionspsychologie. Bd. 1. Bern [u.a.]: Huber.
- MILLER, N./G. MARUYAMA/R. J. BEABER** (1976): Speed of Speech and Persuasion. In: Journal of Personality and Social Psychology 34, Heft 4, S. 615-624.

- MÜLLER, A.** (1999): Die Macht der Stimme. Die Stimme als rhetorischer Wirkungsfaktor: Zur persuasiven Funktion und Wirkung der Prosodie. Bad Iburg: Der Andere Verlag.
- MURRAY, I. R./J. L. ARNOTT** (1993): Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. In: The Journal of the Acoustical Society of America, Bd. 93 (2), S. 1097-1108.
- NEPERT, J.** (1999): Elemente einer akustischen Phonetik. <sup>4</sup>Hamburg: Buske.
- NEUBER, B.** (2002): Prosodische Formen in Funktion. Frankfurt a. M.: Lang.
- NIETZSCHE, F.** (1978): Die Unschuld des Werdens I. Der Nachlass. Stuttgart: Kröner.
- NÖTH, E.** (1991): Prosodische Informationen in der automatischen Spracherkennung. Tübingen: Niemeyer.
- OHALA, J. J.** (1983): Cross-Language Use of Pitch: An Ethological View. In: Phonetica 40, S. 1-18.
- OHALA, J. J.** (1984): An Ethological Perspective on Common Cross-Language Utilization of F<sub>0</sub> of Voice. In: Phonetica 41, S. 1-16.
- OHALA, J. J.** (1994): The frequency code underlies the sound-symbolic use of voice pitch. In: Hinton, J./J. Nichols/J. J. Ohala (Hrsg.): Sound Symbolism. Cambridge: Cambridge University Press, S. 325-347.
- ORTONY, A./T. J. TURNER** (1990): What's Basic about Basic Emotions? In: Psychological Review 97, S. 315-331.
- OSGOOD, C. E.** (1952): The Nature and Measurement of Meaning. In: Psychological Bulletin 3, S. 197-237.
- OSGOOD, C. E./G. J. SUCI/P. H. TANNENBAUM** (1957): The Measurement of Meaning. University of Illinois: Urbana.
- OTTO, J. H./H. A. EULER/H. MANDL** (Hrsg.) (2000): Emotionspsychologie. Weinheim: Beltz.
- PAESCHKE, A./W. F. SENDLMEIER** (2000): Prosodic Characteristics of Emotional Speech: Measurements of Fundamental Frequency Movements. In: Proceedings of the ISCA Workshop on Speech and Emotion. <http://www.qub.ac.uk/en/isca/proceedings>
- PAESCHKE, A.** (2003): Prosodische Analyse emotionaler Sprechweise. Berlin: Logos.
- PAGE, R. A./J. L. BALLOUN** (1978): The Effect of Voice Volume on the Perception of Personality. In: The Journal of Social Psychology 105, S. 65-72.
- PAHN, W.** (1988): Das Registerproblem. In: Gless, D. (Hrsg.): Stimmbildung und Liedgestaltung. Berlin: Volk und Wissen, S. 29-45.

- PARK, C. W./S. M. YOUNG** (1986): Consumer response to television commercials: The impact of involvement and background music on brand attitude formation. *Journal of Marketing Research* 23, S. 11-24.
- PEARCE, W. B./B. J. BROMMEL** (1972): Vocalic Communication in Persuasion. In: *The Quarterly Journal of Speech* 58, Heft 3, S. 298-306.
- PEREIRA, C.** (2000): Dimensions of Emotional Meaning in Speech. In: *Proceedings of the ISCA Workshop on Speech and Emotion*.  
<http://www.qub.ac.uk/en/isca/proceedings>
- PETERSON, R. A./M. P. CANNITO/S. P. BROWN** (1995): An Exploratory Investigation of Voice Characteristics and Selling Effectiveness. In: *Journal of Personal Selling and Sales Management* 15, Heft 1, S. 1-15.
- PETÖFI, J. S.** (1988) (Hrsg.): *Text and Discourse Constitution. Empirical Aspects, Theoretical Approaches*. Berlin [u.a.]: de Gruyter.
- PETTY, R. E./G. L. WELLS/T. C. BROCK** (1976): Disraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification. In: *Journal of Personality and Social Psychology* 34, Heft 5, S. 874-884.
- PETTY, R. E./T. C. BROCK** (1981): Thought disruption and persuasion: Assessing the validity of attitude change experiments. In: Petty, R. E./T. M. Ostrom/T. C. Brock (Hrsg.): *Cognitive Responses in Persuasion*. Hillsdale: Erlbaum, S. 55-79.
- PETTY, R. E./T. M. OSTROM/T. C. BROCK** (1981) (Hrsg.): *Cognitive Responses in Persuasion*. Hillsdale: Erlbaum.
- PETTY, R. E./J. T. CACIOPPO** (1986): *Communication and Persuasion. Central and Periphral Routes to Attitude Change*. New York [u.a.]: Springer.
- PFAU, E.-M./H.-G. STREUBEL** (1982): *Die Behandlung der gestörten Sprechstimme. Stimmfunktionstherapie*. Leipzig: Thieme.
- PHEBY, J.** (1975): *Intonation und Grammatik im Deutschen*. Berlin: Akademie-Verlag.
- PHEBY, J.** (1980): Intonation. In: *Grundzüge einer deutschen Grammatik. Von einem Autorenkollektiv unter der Leitung von Heidolph, K. E./W. Flämig/W. Motsch*. Berlin: Akademie Verlag, S. 839-897.
- PIERREHUMBERT, J.** (1981): Synthesizing intonation. *Journal of the Acoustical Society of America* 70, S. 985-995.
- PLUTCHIK, R.** (Hrsg.) (1980): *Emotion. Theory, Research, and Experience*. Bd.1, *Theories of Emotion*. Orlando [u.a.]: Acad. Press
- POMPINO-MARSCHALL, B.** (1990): *Die Silbenprosodie. Ein elementarer Aspekt der Wahrnehmung von Sprachrhythmus und Sprechtempo*. Tübingen: Niemeyer.

- POMPINO-MARSCHALL, B.** (1995): Einführung in die Phonetik. Berlin [u.a.]: de Gruyter.
- RAITHEL, V./M. HIELSCHER-FASTABEND** (2004): Emotional and Linguistic Perception of Prosody. In: *Folia Phoniatica et Logopaedica* 56, S. 7-13.
- REICHERT, T./S. E. HECKLER/S. JACKSON** (2001): The Effects of Sexual Social Marketing Appeals on Cognitive Processing and Persuasion. In: *The Journal of Advertising* 30, Heft 1, S. 13-27.
- REINKE, K.** (1998): Sprich, damit ich dich sehe. Ein Versuch zum interkulturellen Vergleich der Emotionsrezeption anhand phonetischer Mittel. In: Biege, A./I. Bose (Hrsg.): *Theorie & Empirie in der Sprechwissenschaft*. Hanau [u.a.]: Dausien, S. 188-198.
- REUTER, C.** (2001): Wie und warum in der Geschichte der Klangfarbenforschung meistens am Stimmklang vorbeigeforscht wurde.  
[http://www.uni-koeln.de/phil-fak/muwi/publ/fs\\_fricke/reuter.html](http://www.uni-koeln.de/phil-fak/muwi/publ/fs_fricke/reuter.html)
- ROß, K.** (1994): *Sprecherziehung statt Rhetorik*. Opladen: Westdeutscher Verlag.
- SCHERER, K. R./H. LONDON/J. J. WOLF** (1973): The Voice of Confidence: Paralinguistic Cues and Audience Evaluation. In: *Journal of Research in Personality* 7, S. 31-44.
- SCHERER, K. R.** (1978): Personality inference from voice quality: The loud voice of extroversion. *European Journal of Psychology* 8, S. 467-487.
- SCHERER, K. R.** (1979): Personality markers in speech. In: Scherer, K. R./H. Giles (Hrsg.): *Social markers in speech*. Cambridge: University Press, S. 147-209.
- SCHERER, K. R./H. GILES** (1979) (Hrsg.): *Social markers in speech*. Cambridge: University Press.
- SCHERER, K. R.** (1982): *Vokale Kommunikation. Nonverbale Aspekte des Sprachverhaltens*. Weinheim [u.a.]: Beltz.
- SCHERER, K. R.** (1984): Emotion as a multicomponent process: A model and some cross-cultural data. In: *Review of Personality and Social Psychology* 5, S. 37-63.
- SCHERER, K. R./P. EKMAN** (1984) (Hrsg.): *Approaches to Emotion*. Hillsdale [u.a.]: Erlbaum.
- SCHERER, K. R.** (1986): Vocal Affect Expression: A Review and a Model for Future Research. In: *Psychological Bulletin* 99, Heft 2, S. 143-165.
- SCHERER, K. R./H. G. WALLBOTT** (1990): Ausdruck von Emotionen. In: *Enzyklopädie der Psychologie* 4, C, Bd. 3, S. 345-422.

- SCHERER, K. R./R. BANSE/H. G. WALLBOTT/T. GOLDBECK** (1991): Vocal cues in emotion encoding and decoding. In: *Motivation and Emotion* 15, S. 123-148.
- SCHERER, K. R.** (1996): Emotion. In: Hewston, M./W.Stroebe/G. M. Stephenson (Hrsg.): *Introduction to Social Psychology*. <sup>2</sup>Oxford [u.a.]: Blackwell, S. 279-315.
- SCHERER, K. R.** (2000): Psychological models of emotion. In: Borod, J. (Hrsg.): *The neuropsychology of Emotion*. Oxford [u.a.]: Oxford Univ. Press, S. 137-162.
- SCHERER, K. R.** (2000a): Emotion effects on voice and speech: Paradigms and approaches to evaluation. In: *Proceedings of the ISCA Workshop on Speech and Emotion*.  
<http://www.qub.ac.uk/en/isca/proceedings>
- SCHIERL, T.** (2001): *Text und Bild in der Werbung. Bedingungen, Wirkungen und Anwendungen bei Anzeigen und Plakaten*. Köln: Halem.
- SCHIERL, T.** (2004): *Die Visualisierung des Spots in den Medien*. Köln: Halem.
- SCHIRMER, A./S. A. KOTZ/A. D. FRIEDERICI** (2002): Sex differentiates the role of emotional prosody during word processing. In: *Cognitive Brain Research* 14, S. 228-233.
- SCHIRMER, A./T. STRIANO/A. D. FRIEDERICI** (2005): Sex differences in the preattentive processing of vocal emotional expressions. In: *Cognitive Neuroscience and Neuropsychology* 16 (6), S. 635-639.
- SCHIRMER, A./S. A. KOTZ/A. D. FRIEDERICI** (2005a): On the role of attention for the processing of emotion in speech: Sex differences revisited. In: *Cognitive Brain Research* 24, S. 442-452.
- SCHIRMER, A./S. A. KOTZ** (2006): Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. In: *ScienceDirect* 10 (1), S. 24-30.
- SCHLOSBERG, H.** (1954): Three Dimensions of Emotion. *Psychological Review* 61, S. 81-88.
- SCHMIDT-ATZERT, L.** (1987): Zur umgangssprachlichen Ähnlichkeit von Emotionswörtern. In: *Psychologische Beiträge* 29, S. 140-163.
- SCHMIDT-ATZERT, L.** (1996): *Lehrbuch der Emotionspsychologie*. Stuttgart [u.a.]: Kohlhammer.
- SCHRÖDER, M.** (2000): Experimental Study of Affect Bursts. In: *Proceedings of the ISCA Workshop on Speech and Emotion*.
- SCHRÖDER, M./R.COWIE/E. DOUGLAS-COWIE/M. WESTERDIJK/S. GIELEN** (2001): Acoustic Correlates of Emotion Dimensions in View of Speech Synthesis. In: *Proceedings Eurospeech 1: Aalborg*, S. 87-90.

- SCHRÖDER, M.** (2004): Speech and Emotion Research. An Overview of Research Frameworks and a Dimensional Approach to Emotional Speech Synthesis. In: Barry, W. J. (Hrsg.): Phonus 7, Saarbrücken: Institut für Phonetik Universität des Saarlandes.
- SEIDNER, W./M. BÜTTNER** (1998): Zur Ästhetik der heiseren Singstimme. In: Gundermann, H. (Hrsg.): Die Ausdruckswelt der Stimme. 1. Stuttgarter Stimmtage. Heidelberg: Hüthig, S. 72-76.
- SEIDNER, W./J. WENDLER** (1982): Die Sängerstimme. Phoniatische Grundlagen für die Gesangsausbildung. Berlin: Henschel.
- SELTING, M.** (1995): Prosodie im Gespräch. Tübingen: Niemeyer.
- SHARF, D. J./M. E. LEHMAN** (1984): Relationship between the Speech Characteristics and Effectiveness of Telephone Interviewers. In: Journal of Phonetics 12, S. 219-228.
- SIMPSON, A. P.** (1998): Phonetische Datenbanken des Deutschen in der empirischen Sprachforschung und der phonologischen Theoriebildung. In: Kohler, K. J. (Hrsg.): Arbeitsberichte 33, Kiel: Institut für digitale Sprachverarbeitung.
- SMITH, B. L./B. L. BROWN/W. J. STRONG/A. C. RENCHER** (1975): Effects of Speech Rate on Personality Perception. In: Language and Speech 18, S. 145-152.
- STEWART, M./E. B. RYAN** (1982): Attitudes Toward Younger and Older Adult Speakers: Effects of Varying Speech Rates. In: Journal of Language and Social Psychology 1(2), S. 91-109.
- STOCK, E.** (1980): Untersuchungen zu Form, Bedeutung und Funktion der Intonation im Deutschen. Berlin: Akademie-Verlag.
- STOCK, E./J. SUTTNER** (1991): Wirkungen des Stimm- und Sprechausdrucks. In: Krech, E.-M./G. Richter/E. Stock/J. Suttner (Hrsg.): Sprechwirkung. Grundfragen, Methoden und Ergebnisse ihrer Erforschung. Berlin: Akademie-Verlag, S. 59-142.
- STOCK, E.** (1996): Deutsche Intonation. Leipzig [u.a.]: Langenscheidt.
- STOCK, E.** (1996a): Text und Intonation. In: Sprachwissenschaft 21, S. 211-240.
- STOCK, E.** (1998): Über Sprechrhythmus. In: Kröger, B. J./C. Riek/G. Sachse (Hrsg.): Festschrift für Georg Heike. Frankfurt a. M.: Hector, S. 191-204.
- STROEBE, W./K. JONAS/M. HEWSTONE** (2003) (Hrsg.): Sozialpsychologie. Eine Einführung. <sup>4</sup>Berlin [u.a.]: Springer.
- SUNDBERG, J.** (1997): Die Wissenschaft von der Singstimme. Bonn: Orpheus.

- TAUCHNITZ, J.** (1990): Werbung mit Musik. Theoretische Grundlagen und experimentelle Studien zur Wirkung von Hintergrundmusik in der Rundfunk- und Fernsehwerbung. Heidelberg: Physica.
- TERHARDT, E.** (1998): Akustische Kommunikation. Grundlagen mit Hörbeispielen. Berlin [u.a.]: Springer.
- TEMBROCK, G.** (1977): Tierstimmenforschung. <sup>2</sup>Wittenberg: Ziemsen.
- TEMBROCK, G.** (1996): Akustische Kommunikation bei Säugetieren: Die Stimmen der Säugetiere und ihre Bedeutung. Darmstadt: Wissenschaftliche Buchgesellschaft.
- TILLMANN, H. G.** (1974): Das individuelle Subjekt und seine persönliche Identität im phonetischen Kommunikationsprozess. Hamburg: Buske.
- TISCHER, B.** (1993): Die vokale Kommunikation von Gefühlen. Fortschritte der psychologischen Forschung 18. Weinheim: Beltz.
- TOLKMITT, F. J./K. R. SCHERER** (1986): Effect of Experimentally Induced Stress on Vocal Parameters. In: Journal of Experimental Psychology: Human Perception and Performance 12, Bd. 3, S. 302-313.
- TROJAN, F.** (1952): Der Ausdruck der Sprechstimme: Eine phonetische Lautstilk. <sup>2</sup>Wien [u.a.]: Maudrich.
- TROJAN, F.** (1975): Biophonetik. Mannheim [u.a.]: Bibliographisches Institut.
- TROUVAIN, J.** (2004): Tempo Variation in Speech Production. Implications for Speech Synthesis. In: Barry, W. J. (Hrsg.): Phonus 8, Saarbrücken: Institut für Phonetik Universität des Saarlandes.
- TSURU, S./H. S. FRIES** (1933): A problem in meaning. Journal of general psychology 8, S. 281-284.
- UEDING, G.** (2003): Historisches Wörterbuch der Rhetorik, Bd. 6. Niemeyer: Tübingen.
- ULEMAN, J. S./J. A. BARGH** (1989) (Hrsg.): Unintended thought. New York: Guilford.
- VIEREGGE, W. H.** (1996): Patho-Symbolphonetik: auditive Deskription pathologischer Sprache. Unter Mitarbeit von J. Pahn & H. K. Schutte. In: Zeitschrift für Dialektologie und Linguistik, Beiheft 100. Stuttgart: Steiner.
- WENDLER, J./W. SEIDNER** (1977): Lehrbuch der Phoniatrie. Leipzig: Thieme.
- WENDLER, J./W. SEIDNER/G. KITTEL/U. EYSOLDT** (1996): Lehrbuch der Phoniatrie. Leipzig: Thieme.
- WENGER, M. A./F. N. JONES/M. H. JONES** (1962): Emotional behavior. In: Candland, D. K. (Hrsg.): Emotion: Bodily change. Princeton: Van Nostrand, S. 3-10.

- WILLIAMS, C. E./K. N. STEVENS** (1972): Emotions and Speech: Some Acoustical Correlates. In: The Journal of the Acoustical Society of America 52, Bd. 4 (2), S. 1238-1250.
- WINCKEL, F.** (1960): Die psychoakustische Bewertung des Spektrums. In: Folia Phoniatica 12, S. 129-136.
- WIRTH, G.** (1995): Stimmstörungen. Lehrbuch für Logopäden, Sprachheilpädagogen und Sprechtherapeuten. <sup>4</sup>Köln: Deutscher Ärzte-Verlag.
- WOODALL, W. G./J. K. BURGOON** (1983): Talking Fast and Changing Attitudes: A Critique and Clarification. In: Journal of Nonverbal Behavior 8, Heft 2, S. 126-141.
- WUNDERLICH, D.** (1988): Der Ton macht die Melodie – Zur Phonologie der Intonation des Deutschen. In: Altmann, H. (Hrsg.): Intonationsforschungen. Tübingen: Niemeyer, S. 1-40.
- WUNDT, W.** (1905): Grundzüge der physiologischen Psychologie. Bd. 3. <sup>5</sup>Leipzig: Engelmann.
- WUNDT, W.** (1911): Grundzüge der physiologischen Psychologie. Bd. 3. <sup>6</sup>Leipzig: Kröner.
- WÜSTHOFF, K.** (1999): Die Rolle der Musik in der Film-, Funk- und Fernsehwerbung. <sup>2</sup>Kassel: Merseburger.
- ZENTNER, M. R./K. R. SCHERER** (2000): Partikuläre und integrative Ansätze. In: Otto, J. H./H. A. Euler/H. Mandl (Hrsg.): Emotionspsychologie. Weinheim: Beltz, S. 151-164.
- ZUCKERMANN, M./R. E. DRIVER** (1989): What Sounds beautiful is Good: The Vocal Attractiveness Stereotype: Replication and Elaboration. In: Journal of Nonverbal Behavior 13, S. 67-82.
- ZUCKERMANN, M./H. HODGINS/K. MIYAKE** (1990): The Vocal Attractiveness Stereotype: Replication and Elaboration. In: Journal of Nonverbal Behavior 14, S. 97-112.
- ZWICKER, E.** (1982): Psychoakustik. Berlin: Springer.

## 10 DANKSAGUNG

Ich möchte an dieser Stelle meinen beiden Betreuern Herrn Prof. Dr. Adrian P. Simpson und Herrn PD Dr. Baldur Neuber von der Professur für Phonetik und Sprechwissenschaft am Institut für Germanistische Sprachwissenschaft der FSU-Jena für ihre rege Unterstützung im Entstehungsprozess dieser Arbeit danken. Ihre kritischen Anmerkungen und ihre konstruktiven Vorschläge haben mich durch alle Arbeitsphasen hindurch, von der Themenfindung bis zum fertigen Manuskript, begleitet. Dabei stellte das Schaffen einer Arbeitsatmosphäre, in der eigene Ideen und Vorschläge aufgegriffen und im Gespräch diskutiert wurden, eine unverzichtbare Basis für die anhaltende Motivation über die gesamte Arbeitszeit dar.

In anregenden Gesprächen haben mich meine Kolleginnen Frau Dr. Beate Rues, Frau Romy Baumgarten und Frau Ramona Benkenstein, immer wieder durch Zuspruch in meinem Arbeiten bestärkt, und mir dadurch die Sicherheit gegeben, auf dem richtigen Weg zu sein. Auch die regelmäßigen Arbeitstage in der Bibliothek mit Claudia Ernst waren ein unersetzlicher Gewinn für das Entstehen der Arbeit. Dafür sei allen genannten Personen herzlichst gedankt.

In einer experimentellen Arbeit ist man nicht selten auf die Hilfe von Fachleuten anderer Disziplinen angewiesen. Ich bin Herrn Thoralf Müller (Mediendesigner) und Herrn Michael Donatus für die technische Erstellung des Werbevideos in unendlicher Kleinarbeit, und Herrn Andreas Wolf für die Betreuung der gesamten Statistik zu größtem Dank verpflichtet, die sich mit großer Geduld und Kreativität in das experimentelle Design der Arbeit eingedacht haben.

Diese Arbeit hätte in dieser Form kein Ende gefunden, wenn nicht die Nähe und Zuwendung meines Mannes Christoph, Tag für Tag ein treuer Begleiter gewesen wäre. Seine unendliche Liebe und Zuversicht, aber auch sein gelegentlich sanfter Druck, haben mir den Rücken gestärkt und mich an vermeintlichen Tiefpunkten wieder aufgebaut. Dafür danke ich ihm zutiefst.

# 11      **TABELLARISCHER LEBENS LAUF**

## **PERSÖNLICHE DATEN**

Name	Beate Redecker, geb. Diener
Geburtstag	13. Januar 1976
Geburtsort	Stollberg/Erzgeb.
Familienstand	verheiratet
Anschrift	Nollendorfer Str. 19 07743 Jena

## **SCHULAUSSBILDUNG**

08/1982 – 07/1990	Besuch der POS „J. R. Becher“ in Dorndorf/S.
08/1990 – 07/1994	Besuch des Gymnasiums „Adolf Reichwein“ in Jena Abschluss: Allgemeine Hochschulreife

## **AUSLANDSAUFENTHALT**

09/1994 – 07/1995	Au-pair-Aufenthalt in Versailles (Frankreich) Sprachkurs an der Alliance Française (Zertifikat)
-------------------	--

## **HOCHSCHULAUSSBILDUNG**

WS 1995/1996	Immatrikulation an der FSU Jena in den Fächern Romanistik, DaF, Phonetik/Sprechwissenschaft
WS 1996/1997	Fachrichtungswechsel von Romanistik zu Erziehungswissenschaften
SS 1997	Magisterzwischenprüfung in den Fächern DaF (1,3) und Phonetik/Sprechwissenschaft (1,7)
WS 1997/1998	Magisterzwischenprüfung im Fach Erziehungswissenschaften (1,0)
08/2000 – 02/2001	Magisterarbeit zum Thema: Untersuchung prosodischer Aspekte des Emotionsausdrucks unter Berücksichtigung ihrer Bedeutung beim kindlichen Spracherwerb. (Note: 1,0)
SS 2001	Magisterprüfungen: Erziehungswissenschaften (1,0), DaF (1,0), Phonetik/Sprechwissenschaft (1,3)
09/2001	Studienabschluss M.A. mit der Gesamtnote „sehr gut“ (1,0)

## STUDIENAUFENTHALTE

- WS 1998/1999                      Aufenthalt in Oxford (GB) an der Oxford Brookes University als Assistent Teacher
- 04/2001                              Aufenthalt in Tampere (Finnland) an der Multi Media School. Dreharbeiten im Rahmen eines EU-Projektes zur Erstellung eines Deutschlehrwerkes auf CD-Rom.

## SONSTIGE AKTIVITÄTEN

- WS 1997/1998 –  
WS 2000/2001                      Studentische Hilfskraft am Lehrstuhl für Phonetik/Sprechwissenschaft bei Prof. Dr. G. Meinhold
- WS 2000/2001                      Studentisches Mitglied der Berufungskommission für die Nachfolgeprofessur Phonetik/ Sprechwissenschaft an der FSU Jena

## PRAKTIKA & TUTORIEN

- SS 1998                              Unterrichtspraktische Übungen: Unterrichten japanischer Studenten unter Beisitzen eines Dozenten
- WS 1998/1999                      Assistent Teacher an der Oxford Brookes University in Oxford (GB): Leitung eines Deutschkurses mit dem Schwerpunkt *Konversation*
- SS 1999 – SS 2001                Tutorium am Institut DaF/DaZ der FSU Jena: Unterrichten japanischer Studenten im Fach *korrektive Phonetik* (Anfänger & Fortgeschrittene)
- Sommer 2002                      Sprachassistentin an der Sommeruniversität der Open University (GB): Konversationskurse bei fortgeschrittenen Lernern
- 02/2000 – 03/2000                Praktikum am LKA Brandenburg, Sachgebiet Forensische Phonetik und Sprechererkennung

## TÄTIGKEIT

- Seit 01/10/2001                      Wissenschaftliche Mitarbeiterin am Lehrstuhl für Phonetik/Sprechwissenschaft der FSU Jena (Institut für Germanistische Sprachwissenschaft/Philosophische Fakultät)
- Lehrtätigkeit                        \* Sprechererziehung (Lehramtsstudenten)  
    \* Sprechbildung I und II (Fachstudenten)  
    \* Programmgestaltung (Fachstudenten)  
    \* Rhetorik (Fachstudenten und Hörer aller Fakultäten)  
    \* Einführung in die Phonetik/Phonologie (Germanisten)

## **12 EHRENWÖRTLICHE ERKLÄRUNG**

Ich erkläre hiermit, dass mir die Promotionsordnung der Philosophischen Fakultät der Friedrich-Schiller-Universität Jena bekannt ist.

Ferner erkläre ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung des Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise unentgeltlich geholfen:

1. Herr Adrian P. Simpson (Einsprechen der Sprechfassungen)
2. Herr Thoralf Müller (Erstellen des Werbevideos)
3. Herr Andreas Wolf (statistische Beratung)
4. Frau Romy Baumgarten, Frau Ramona Benkenstein, Herr Frank Lorenz, Herr Baldur Neuber und Frau Beate Rues (Kontrollhörer in der auditiven Analyse)

Weitere Personen waren an der inhaltlich-materiellen Erstellung der Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Arbeit stehen.

Die Arbeit wurde bisher weder im Inland noch Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

## **ANHANG**

A.1	Fragebogen zur Wirkungsanalyse	172
A.2	Fragebogen zur auditiven Analyse	175
A.3	Statistik (Rohdaten)	177
A.4	CD mit den untersuchten Werbespots	193

## FRAGEBOGEN

Sie sehen jetzt eine Parfumwerbung.

Entscheiden Sie, inwiefern Sie die folgenden Eigenschaften (a) mit dem Werbespot, (b) mit der Person, und (c) mit dem Duft verbinden.

**Es gibt keine richtigen und falschen Antworten. Entscheiden Sie sich im Zweifelsfall für die Variante, der sie spontan am ehesten zustimmen würden.**

### A: WIE WIRKT DER WERBESPOT AUF SIE?

EIGENSCHAFT	Trifft gar nicht zu -2	-1	0	+1	Trifft völlig zu +2
exklusiv					
edel					
farblos					
professionell					
schlicht					
preiswert					
animierend					
gewöhnlich					

Weitere Eigenschaften: \_\_\_\_\_

### B: WIE WIRKT DIE PERSON AUF SIE?

EIGENSCHAFT	Trifft gar nicht zu -2	-1	0	+1	Trifft völlig zu +2
charmant					
dynamisch					
erotisch					
weich					
aufdringlich					
charismatisch					
männlich					
abstoßend					

Weitere Eigenschaften: \_\_\_\_\_

**C: WIE BEURTEILEN SIE DEN DUFT?**

EIGENSCHAFT	Trifft gar nicht zu -2	-1	0	+1	Trifft völlig zu +2
blumig					
moschusartig					
lieblich					
markant					
süß					
tabakartig					
dezent					
warm					
herb					
fruchtig					
maskulin					
leicht					
elegant					
frisch					
holzig					
sportlich					
würzig					
spritzig					
klassisch					
kühl					
vitalisierend					
schwer					

Weitere Eigenschaften: \_\_\_\_\_

Wann würden Sie das Parfüm tragen:  im Alltag  in der Freizeit  
 am Abend  nie

## ALLGEMEINE FRAGEN

- Ist Ihnen das Parfüm bekannt?  ja  nein
- Wenn Sie ♂ sind:  
Würden Sie das Parfüm testen?  ja  nein
- Wenn Sie ♀ sind:  
Würden Sie das Parfüm ihrem  
Partner kaufen oder schenken?  ja  nein

## ANGABEN ZU IHRER PERSON:

Alter  unter 20  21-24  25-28  29+

Geschlecht  männlich  weiblich

Studiengang (Fakultät)  naturwissenschaftlich  
 geisteswissenschaftlich  
 sozialwissenschaftlich  
 anderer, \_\_\_\_\_

Wie oft benutzen Sie Parfüm?  täglich  mehrmals pro Woche  einmal pro Woche  
 weniger als einmal pro Woche  nie

Spielen Gerüche/Düfte für Sie eine Rolle?  gar keine Rolle 

-2	-1	0	+1	+2
----	----	---	----	----

 sehr große Rolle

Halten Sie sich eher für einen  akustischen Typ  
 visuellen Typ  
 olfaktorischen Typ  
 keinen besonderen Typ

Wo sind Sie aufgewachsen?  Großstadt (100.000+)  Kleinstadt (5.000 - 100.000)  Dorf (≤ 5.000)

**FRAGEBOGEN ZUR BEURTEILUNG DER STIMME (FORMKRITERIEN)****1. HABITUELLE STIMMQUALITÄTEN**

(bezogen auf die physiologischen Voraussetzungen des Sprechers)

**SPRECHSTIMMLAGE**

tiefe Lage                       mittlere Lage                       hohe Lage

**INTENSITÄT**

geringe Intensität                       mittlere Intensität                       hohe Intensität

**SPRECHGESCHWINDIGKEIT**

langsam                       mittel                       schnell

**STIMMEINSÄTZE UND STIMMQUALITÄT**

klar                       be-/verhaucht                       knarrend

**KLANGFÜLLE**

resonanzreich  
(kräftig/voll)                       mittel                       resonanzarm  
(zart/dünn)

**KLANGFARBE**

hell                       mittel                       dunkel

weich                       mittel                       hart

warm                       mittel                       kalt



## Faktor 1: Werbespot positiv

Mehrfachvergleiche  
Abhängige Variable: **exklusiv**  
Tukey-HSD

		Mittlere Dif- ferenz (I-J)	STAW	Signifikanz
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.3088	0.21357	0.320
	markant	0.1051	0.26242	0.916
schwer	leicht	0.3088	0.21357	0.320
	markant	0.4139	0.24225	0.205
markant	leicht	-0.1051	0.26242	0.916
	schwer	-0.4139	0.24225	0.205

Mehrfachvergleiche  
Abhängige Variable: **edel**  
Tukey-HSD

		Mittlere Dif- ferenz (I-J)	STAW	Signifikanz
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.5658*	0.21306	0.024*
	markant	0.4391	0.25943	0.211
schwer	leicht	0.5658*	0.21306	0.024*
	markant	1.0049**	0.23911	0.000**
markant	leicht	-0.4391	0.25943	0.211
	schwer	-1.0049**	0.23911	0.000**

Mehrfachvergleiche  
Abhängige Variable: **professionell**  
Tukey-HSD

		Mittlere Dif- ferenz (I-J)	STAW	Signifikanz
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.5322*	0.19746	0.021*
	markant	0.1357	0.24053	0.839
schwer	leicht	0.5322*	0.19746	0.021*
	markant	0.6679**	0.22025	0.008**
markant	leicht	-0.1357	0.24053	0.839
	schwer	-0.6679**	0.22025	0.008**

Basiert auf beobachteten Mittelwerten.

\* Die mittlere Differenz ist auf der Stufe .05 signifikant.

\*\* Die mittlere Differenz ist auf der Stufe .01 hoch signifikant

Mehrfachvergleiche  
Abhängige Variable: **animierend**  
Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.5336*	0.22255	0.046*
	markant	0.2215	0.26863	0.688
schwer	leicht	0.5336*	0.22255	0.046*
	markant	0.7552**	0.24720	0.007**
markant	leicht	-0.2215	0.26863	0.688
	schwer	-0.7552**	0.24720	0.007**

Mehrfachvergleiche  
Abhängige Variable: **Wirkung Werbespot positiv (exklusiv, edel, professionell, animierend)**  
Tamhane

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.4959**	0.13738	0.001**
	markant	0.2344	0.19686	0.559
schwer	leicht	0.4959**	0.13738	0.001**
	markant	0.7303**	0.18781	0.001**
markant	leicht	-0.2344	0.19686	0.559
	schwer	-0.7303**	0.18781	0.001**

## Faktor 2: Werbespot negativ

Mehrfachvergleiche  
Abhängige Variable: **schlicht**  
Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.4243	0.24460	0.196
	markant	0.1481	0.30065	0.875
schwer	leicht	-0.4243	0.24460	0.196
	markant	-0.2762	0.27578	0.577
markant	leicht	-0.1481	0.30065	0.875
	schwer	0.2762	0.27578	0.577

Mehrfachvergleiche  
Abhängige Variable: **preiswert**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.0533	0.20114	0.962
	markant	-0.2059	0.24291	0.674
schwer	leicht	0.0533	0.20114	0.962
	markant	-0.1525	0.22206	0.771
markant	leicht	0.2059	0.24291	0.674
	schwer	0.1525	0.22206	0.771

Mehrfachvergleiche  
Abhängige Variable: **gewöhnlich**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.3611	0.23527	0.277
	markant	-0.1608	0.28472	0.839
schwer	leicht	-0.3611	0.23527	0.277
	markant	-0.5220	0.26146	0.117
markant	leicht	0.1608	0.28472	0.839
	schwer	0.5220	0.26146	0.117

Mehrfachvergleiche  
Abhängige Variable: **Wirkung Werbespot negativ (schlicht, preiswert, gewöhnlich)**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.2632	0.17325	0.285
	markant	-0.0458	0.20967	0.974
schwer	leicht	-0.2632	0.17325	0.285
	markant	-0.3090	0.19254	0.247
markant	leicht	0.0458	0.20967	0.974
	schwer	0.3090	0.19254	0.247

Mehrfachvergleiche  
 Abhängige Variable: **Wirkung Werbespot gesamt (positiv & negativ)**  
 Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.3964**	0.13049	0.008**
	markant	0.1471	0.15792	0.621
schwer	leicht	0.3964**	0.13049	0.008**
	markant	0.5435**	0.14502	0.001**
markant	leicht	-0.1471	0.15792	0.621
	schwer	-0.5435**	0.14502	0.001**

**Anmerkung: Item „farblos“ fällt aus beiden Faktoren (Werbespot<sub>pos</sub> und Werbespot<sub>neg</sub>) heraus → möglicherweise wegen der schwarz-weiß-Darstellung des Videos**

### Faktor 1: Person positiv

Mehrfachvergleiche  
 Abhängige Variable: **charmant**  
 Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.6474**	0.21949	0.010**
	markant	-0.2037	0.26795	0.728
schwer	leicht	0.6474**	0.21949	0.010**
	markant	0.4436	0.24646	0.173
markant	leicht	0.2037	0.26795	0.728
	schwer	-0.4436	0.24646	0.173

Mehrfachvergleiche  
 Abhängige Variable: **dynamisch**  
 Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.2174	0.19445	0.505
	markant	0.0856	0.23687	0.931
schwer	leicht	0.2174	0.19445	0.505
	markant	0.3030	0.21690	0.345
markant	leicht	-0.0856	0.23687	0.931
	schwer	-0.3030	0.21690	0.345

Mehrfachvergleiche  
Abhängige Variable: **erotisch**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.7786**	0.20060	0.000**
	markant	-0.0652	0.24225	0.961
schwer	leicht	0.7786**	0.20060	0.000**
	markant	0.7133**	0.22145	0.004**
markant	leicht	0.0652	0.24225	0.961
	schwer	-0.7133**	0.22145	0.004**

Mehrfachvergleiche  
Abhängige Variable: **charismatisch**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.1745	0.19964	0.657
	markant	0.1925	0.24006	0.702
schwer	leicht	0.1745	0.19964	0.657
	markant	0.3671	0.21736	0.213
markant	leicht	-0.1925	0.24006	0.702
	schwer	-0.3671	0.21736	0.213

Mehrfachvergleiche  
Abhängige Variable: **männlich**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.5384**	0.16398	0.004**
	markant	-0.3498	0.19844	0.186
schwer	leicht	0.5384**	0.16398	0.004**
	markant	0.1886	0.18223	0.556
markant	leicht	0.3498	0.19844	0.186
	schwer	-0.1886	0.18223	0.556

Mehrfachvergleiche

Abhängige Variable: **Wirkung Person positiv (charmant, dynamisch, erotisch, charismatisch, männlich)**

Tamhane

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.5177**	0.13030	0.001**
	markant	-0.1077	0.18055	0.910
schwer	leicht	0.5177**	0.13030	0.001**
	markant	0.4100*	0.15134	0.028*
markant	leicht	0.1077	0.18055	0.910
	schwer	-0.4100*	0.15134	0.028*

## Faktor 2: Person negativ

Mehrfachvergleiche

Abhängige Variable: **weich**

Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.0945	0.21694	0.901
	markant	0.6522*	0.26426	0.039*
schwer	leicht	0.0945	0.21694	0.901
	markant	0.7467**	0.24198	0.007**
markant	leicht	-0.6522*	0.26426	0.039*
	schwer	-0.7467**	0.24198	0.007**

Mehrfachvergleiche

Abhängige Variable: **aufdringlich**

Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.5730**	0.19548	0.011**
	markant	-0.3270	0.24554	0.380
schwer	leicht	-0.5730**	0.19548	0.011**
	markant	-0.9000**	0.22698	0.000**
markant	leicht	0.3270	0.24554	0.380
	schwer	0.9000**	0.22698	0.000**

Mehrfachvergleiche  
Abhängige Variable: **abstoßend**  
Tukey-HSD

		Mittlere Dif- ferenz (I-J)	STAW	Signifikanz
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.3106	0.13587	0.061
	markant	-0.1560	0.16587	0.615
schwer	leicht	-0.3106	0.13587	0.061
	markant	-0.4667**	0.15256	0.007**
markant	leicht	0.1560	0.16587	0.615
	schwer	0.4667**	0.15256	0.007**

Mehrfachvergleiche  
Abhängige Variable: **Wirkung Person negativ (weich, aufdringlich, abstoßend)**  
Tamhane

		Mittlere Dif- ferenz (I-J)	STAW	Signifikanz
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.3342	0.14196	0.062
	markant	-0.3924	0.18333	0.103
schwer	leicht	-0.3342	0.14196	0.062
	markant	-0.7267**	0.14962	0.000**
markant	leicht	0.3924	0.18333	0.103
	schwer	0.7267**	0.14962	0.000**

## Faktor 1: Duft lieblich-fruchtig

Mehrfachvergleiche  
Abhängige Variable: **blumig**  
Tukey-HSD

		Mittlere Dif- ferenz (I-J)	STAW	Signifikanz
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.0693	0.13441	0.864
	markant	-0.1586	0.16232	0.593
schwer	leicht	-0.0693	0.13441	0.864
	markant	-0.2278	0.14838	0.277
markant	leicht	0.1586	0.16232	0.593
	schwer	0.2278	0.14838	0.277

Mehrfachvergleiche  
Abhängige Variable: **lieblich**  
Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.1365	0.18989	0.753
	markant	-0.2444	0.23131	0.543
schwer	leicht	0.1365	0.18989	0.753
	markant	-0.1079	0.21181	0.867
markant	leicht	0.2444	0.23131	0.543
	schwer	0.1079	0.21181	0.867

Mehrfachvergleiche  
Abhängige Variable: **süß**  
Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.0388	0.18368	0.976
	markant	-0.3478	0.22182	0.263
schwer	leicht	-0.0388	0.18368	0.976
	markant	-0.3867	0.20278	0.140
markant	leicht	0.3478	0.22182	0.263
	schwer	0.3867	0.20278	0.140

Mehrfachvergleiche  
Abhängige Variable: **dezent**  
Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.1467	0.22192	0.786
	markant	0.0485	0.26973	0.982
schwer	leicht	0.1467	0.22192	0.786
	markant	0.1952	0.24585	0.707
markant	leicht	-0.0485	0.26973	0.982
	schwer	-0.1952	0.24585	0.707

Mehrfachvergleiche  
Abhängige Variable: **warm**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.1580	0.21486	0.743
	markant	-0.0678	0.25947	0.963
schwer	leicht	0.1580	0.21486	0.743
	markant	0.0902	0.23720	0.923
markant	leicht	0.0678	0.25947	0.963
	schwer	-0.0902	0.23720	0.923

Mehrfachvergleiche  
Abhängige Variable: **fruchtig**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.0586	0.16391	0.932
	markant	0.1240	0.19794	0.806
schwer	leicht	-0.0586	0.16391	0.932
	markant	0.0655	0.18095	0.930
markant	leicht	-0.1240	0.19794	0.806
	schwer	-0.0655	0.18095	0.930

Mehrfachvergleiche  
Abhängige Variable: **leicht**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.1330	0.21338	0.807
	markant	0.2928	0.25769	0.493
schwer	leicht	0.1330	0.21338	0.807
	markant	0.4259	0.23557	0.170
markant	leicht	-0.2928	0.25769	0.493
	schwer	-0.4259	0.23557	0.170

Mehrfachvergleiche

Abhängige Variable: **Duft lieblich/fruchtig (blumig, lieblich, süß, dezent, warm, fruchtig, leicht)**

Tamhane

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.0609	0.11140	0.929
	markant	-0.0527	0.17345	0.987
schwer	leicht	0.0609	0.11140	0.929
	markant	0.0082	0.16117	10.000
markant	leicht	0.0527	0.17345	0.987
	schwer	-0.0082	0.16117	1.000

## Faktor 2: Duft frisch-vitalisierend-sportlich

Mehrfachvergleiche

Abhängige Variable: **maskulin**

Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>Standardfehler</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.2962	0.15388	0.135
	markant	-0.1637	0.18583	0.653
schwer	leicht	0.2962	0.15388	0.135
	markant	0.1325	0.16988	0.716
markant	leicht	0.1637	0.18583	0.653
	schwer	-0.1325	0.16988	0.716

Mehrfachvergleiche

Abhängige Variable: **elegant**

Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.6035**	0.20116	0.009**
	markant	-0.0141	0.24293	0.998
schwer	leicht	0.6035**	0.20116	0.009**
	markant	0.5894*	0.22207	0.024*
markant	leicht	0.0141	0.24293	0.998
	schwer	-0.5894*	0.22207	0.024*

Mehrfachvergleiche  
Abhängige Variable: **frisch**  
Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.0400	0.18875	0.976
	markant	0.1176	0.22795	0.864
schwer	leicht	0.0400	0.18875	0.976
	markant	0.1576	0.20838	0.730
markant	leicht	-0.1176	0.22795	0.864
	schwer	-0.1576	0.20838	0.730

Mehrfachvergleiche  
Abhängige Variable: **sportlich**  
Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.1594	0.16417	0.596
	markant	0.0563	0.19826	0.957
schwer	leicht	0.1594	0.16417	0.596
	markant	0.2157	0.18124	0.461
markant	leicht	-0.0563	0.19826	0.957
	schwer	-0.2157	0.18124	0.461

Mehrfachvergleiche  
Abhängige Variable: **spritzig**  
Tukey-HSD

		<b>Mittlere Diferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.1439	0.21602	0.783
	markant	0.0852	0.26429	0.944
schwer	leicht	0.1439	0.21602	0.783
	markant	0.2292	0.24019	0.607
markant	leicht	-0.0852	0.26429	0.944
	schwer	-0.2292	0.24019	0.607

Mehrfachvergleiche  
Abhängige Variable: **kühl**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.1243	0.20162	0.811
	markant	0.0691	0.24349	0.957
schwer	leicht	-0.1243	0.20162	0.811
	markant	-0.0553	0.22258	0.967
markant	leicht	-0.0691	0.24349	0.957
	schwer	0.0553	0.22258	0.967

Mehrfachvergleiche  
Abhängige Variable: **vitalisierend**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.2046	0.21101	0.597
	markant	-0.1266	0.25482	0.873
schwer	leicht	0.2046	0.21101	0.597
	markant	0.0780	0.23295	0.940
markant	leicht	0.1266	0.25482	0.873
	schwer	-0.0780	0.23295	0.940

Mehrfachvergleiche  
Abhängige Variable: **Duft Faktor frisch/vitalisierend/sportlich (maskulin, leicht, elegant, frisch, sportlich, spritzig, kühl, vitalisierend)**  
Tamhane

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.1802	0.10475	0.243
	markant	0.0325	0.16127	0.996
schwer	leicht	0.1802	0.10475	0.243
	markant	0.2128	0.15352	0.433
markant	leicht	-0.0325	0.16127	0.996
	schwer	-0.2128	0.15352	0.433

### Faktor 3: Duft würzig-holzig-herb

Mehrfachvergleiche  
Abhängige Variable: **moschusartig**  
Tukey-HSD

		Mittlere Diferenz (I-J)	STAW	Signifikanz
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.2316	0.24718	0.618
	markant	-0.2303	0.30447	0.730
schwer	leicht	-0.2316	0.24718	0.618
	markant	-0.4619	0.27592	0.219
markant	leicht	0.2303	0.30447	0.730
	schwer	0.4619	0.27592	0.219

Mehrfachvergleiche  
Abhängige Variable: **markant**  
Tukey-HSD

		Mittlere Diferenz (I-J)	STAW	Signifikanz
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.1689	0.15811	0.535
	markant	0.0141	0.19218	0.997
schwer	leicht	0.1689	0.15811	0.535
	markant	0.1830	0.17516	0.550
markant	leicht	-0.0141	0.19218	0.997
	schwer	-0.1830	0.17516	0.550

Mehrfachvergleiche  
Abhängige Variable: **tabakartig**  
Tukey-HSD

		Mittlere Diferenz (I-J)	STAW	Signifikanz
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.0258	0.22973	0.993
	markant	-0.4079	0.27743	0.308
schwer	leicht	-0.0258	0.22973	0.993
	markant	-0.4337	0.25361	0.205
markant	leicht	0.4079	0.27743	0.308
	schwer	0.4337	0.25361	0.205

Mehrfachvergleiche  
Abhängige Variable: **herb**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.1583	0.21216	0.737
	markant	-0.3845	0.26078	0.306
schwer	leicht	0.1583	0.21216	0.737
	markant	-0.2263	0.23921	0.612
markant	leicht	0.3845	0.26078	0.306
	schwer	0.2263	0.23921	0.612

Mehrfachvergleiche  
Abhängige Variable: **maskulin**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.2962	0.15388	0.135
	markant	-0.1637	0.18583	0.653
schwer	leicht	0.2962	0.15388	0.135
	markant	0.1325	0.16988	0.716
markant	leicht	0.1637	0.18583	0.653
	schwer	-0.1325	0.16988	0.716

Mehrfachvergleiche  
Abhängige Variable: **holzig**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.1698	0.20115	0.676
	markant	-0.2046	0.24229	0.676
schwer	leicht	-0.1698	0.20115	0.676
	markant	-0.3744	0.22196	0.214
markant	leicht	0.2046	0.24229	0.676
	schwer	0.3744	0.22196	0.214

Mehrfachvergleiche  
Abhängige Variable: **würzig**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.1058	0.24065	0.899
	markant	-0.0064	0.28988	1.000
schwer	leicht	0.1058	0.24065	0.899
	markant	0.0994	0.26555	0.926
markant	leicht	0.0064	0.28988	1.000
	schwer	-0.0994	0.26555	0.926

Mehrfachvergleiche  
Abhängige Variable: **klassisch**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.1032	0.20795	0.873
	markant	0.2788	0.25114	0.509
schwer	leicht	0.1032	0.20795	0.873
	markant	0.3820	0.22958	0.223
markant	leicht	-0.2788	0.25114	0.509
	schwer	-0.3820	0.22958	0.223

Mehrfachvergleiche  
Abhängige Variable: **schwer**  
Tukey-HSD

		<b>Mittlere Dif- ferenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	0.0988	0.23834	0.910
	markant	-0.3772	0.28783	0.391
schwer	leicht	-0.0988	0.23834	0.910
	markant	-0.4761	0.26312	0.170
markant	leicht	0.3772	0.28783	0.391
	schwer	0.4761	0.26312	0.170

Mehrfachvergleiche

Abhängige Variable: **Duft Faktor würzig/holzig/herb (moschusartig, markant, tabakartig, herb, maskulin, holzig, würzig, klassisch, schwer)**

Tamhane

		<b>Mittlere Differenz (I-J)</b>	<b>STAW</b>	<b>Signifikanz</b>
<b>(I) Gruppe</b>	<b>(J) Gruppe</b>			
leicht	schwer	-0.0322	0.12870	0.992
	markant	-0.1414	0.17295	0.801
schwer	leicht	0.0322	0.12870	0.992
	markant	-0.1092	0.14613	0.841
markant	leicht	0.1414	0.17295	0.801
	schwer	0.1092	0.14613	0.841