

Ines Bose

Stimmlich-artikulatorischer Ausdruck und Sprache

Abstract

Seit einigen Jahren befassen sich zahlreiche geisteswissenschaftliche Arbeiten verstärkt mit der Stimme in ihrer Bedeutung für die menschliche Kommunikation. Aufgrund der Vielschichtigkeit und Ambivalenz des Phänomens wird oft von einem sehr weiten, eher metaphorischen Begriff von Stimme ausgegangen. In der Sprechwissenschaft, die traditionell einen vor allem empirischen und didaktischen Zugriff auf die Sprechstimme hat, wird dagegen mit einem vergleichsweise engen, physiologischen Begriff von Stimme operiert, im Sinne einer Körperfunktion, als Muskelaktivitätsmuster unterschiedlicher Ausprägung, Gestalt und Funktion. In engem Bezug zur klinischen Sprechstimmdiagnostik und Phoniatrie wird Stimme betrachtet als Organ, dessen Anatomie und Physiologie zu beschreiben ist. In engem Bezug zur Phonetik werden Stimmgebung und -wirkung, stimmlich-artikulatorische Ausdrucksformen merkmalsanalytisch auditiv und akustisch beschrieben. In engem Bezug zur Linguistik, Rhetorik und Soziophonetik wird Stimme betrachtet als Resultat der Stimmgebung; Gegenstand sind Sprechwirkung und interaktive Ausarbeitung der Stimme in ihrer Verwobenheit mit sprachlichen und körperlichen Ausdrucksformen. Hierbei wird die Stimme als Bestandteil des stimmlich-artikulatorischen Ausdrucks aufgefasst, insofern als Teil von persönlicher und sozialer Identität, als Trägerin von ästhetisch-künstlerischem und emotionalem Ausdruck.

Methoden zur empirischen Beschreibung des sprechstimmlichen Ausdrucks werden exemplarisch vorgestellt, sowohl anhand von Arbeiten, in denen stimmlich-artikulatorische Merkmale beschrieben und klassifiziert werden (Stimmphysiologie-, Emotionsforschung), als auch anhand von Arbeiten, in denen der stimmlich-artikulatorische Ausdruck in seiner Wirkung auf Hörer (Sprechwirkungsforschung) und in seiner interaktiven Ausarbeitung (Gesprächsforschung) betrachtet wird. Aus den Ergebnissen der empirischen Studien wird deutlich, inwieweit stimmlich-artikulatorische Ausdrucksformen als Bedeutungsträger fungieren und zur Verdeutlichung der Verständigung beitragen können.

1. Einleitung

Eine Stimme ruft bei Hörern zunächst einen Gesamteindruck hervor, aufgrund dessen sie Zuordnungen treffen, z.B. zur geografischen Herkunft, zur Nationalität oder zur Stimmung des Sprechers. Sie tun das auch dann, wenn sie den verbalen Gehalt nicht verstehen, und sogar dann, wenn die Hinweise aus dem Gehörtem widersprüchlich sind. Diese Alltagserfahrung hat bereits Edgar Allan Poe vor mehr als 150 Jahren in seiner Kurzgeschichte „Die Morde in der Rue Morgue“ sehr treffend und eindrucksvoll verarbeitet: Nach einem vollkommen unerklärlichen Doppelmord an zwei Pariser Frauen sagen zahlreiche Ohrenzeugen aus, dass sie im Haus der Opfer einen Streit gehört hätten. Neben den ihnen bekannten Stimmen der bei-

den Frauen hätten sie noch eine weitere, unbekannte Stimme wahrgenommen. Diese schrille Stimme deuten die Zeugen allerdings sehr verschieden, und der geniale Detektiv Auguste Dupin stutzt:

Doch was die schrille Stimme angeht, ist das Eigenartige nicht, dass sie uneins waren – sondern dass, obwohl ein Italiener, ein Engländer, ein Spanier, ein Holländer und ein Franzose sie zu beschreiben versuchten, jeder von der Stimme eines Ausländers sprach. Jeder ist sich sicher, dass es nicht die Stimme eines seiner Landsleute war. Mehr noch, ein jeder vergleicht sie, nicht mit der Stimme eines Sprechers einer Nation, mit deren Sprache er vertraut ist, sondern im Gegenteil. Der Franzose hält sie für die Stimme eines Spaniers, und ‚hätte einzelne Wörter ausmachen können, wäre er des Spanischen mächtig‘. Ein Holländer beharrt darauf, dass es die Stimme eines Franzosen gewesen sei, doch entnehmen wir der Aussage, dass der Zeuge, ‚der kein Französisch spreche, von einem Übersetzer befragt wurde‘. Der Engländer hält die Stimme für die eines Deutschen, und ‚versteh kein Deutsch‘. Der Spanier ‚ist sich sicher‘, dass es ein Engländer war, doch er urteilt ‚nur nach der Betonung‘, da ‚er über keine Englischkenntnisse verfüge‘. Der Italiener glaubt, es war ein Russe, ‚doch er habe nie mit einem gebürtigen Russen gesprochen‘. Ein zweiter Franzose ist anderer Meinung als der erste, und ist sich sicher, dass es die Stimme eines Italieners war; er habe sich, ‚da er jener Sprache nicht mächtig sei‘, wie der Spanier ‚von der Betonung überzeugen lassen‘. Wie überhaupt eigenartig muss die Stimme in Wirklichkeit gewesen sein, um solche Zeugenaussagen hervorzurufen! – in deren Klängen selbst die Bewohner von fünf großen Teilen Europas nichts Vertrautes erkennen konnten! (Poe 1841 [2009], S. 35 f.)

Dupin kommt es also höchst eigentümlich vor, dass die Aussagen so stark voneinander abweichen. Als guter Feldforscher unterstellt der Detektiv den Zeugen nicht generelle Unfähigkeit zu präziser Beschreibung bzw. Zuordnung, sondern vermutet wichtige Gründe im stimmlichen Signal. So aufmerksam geworden, findet er im Hause der Ermordeten weitere höchst ungewöhnliche Spuren und kann schließlich den eigentlichen Täter identifizieren: einen Orang Utan.

Rezipienten deuten eine konkrete Stimme immer vor dem Hintergrund ihrer Erfahrungen bzw. Erwartungen als Ausdruck von Persönlichkeitseigenschaften, Emotionen, Einstellungen und Interaktionsrollen. Für diese Deutungen greifen sie auf verschiedene Signale gleichzeitig zu, im oben angeführten Beispiel neben dem Stimmklang z.B. auch auf Akzentuierung und Rhythmus, und sie können das Gehörte gezielt analysieren, sich auf bestimmte Eigenschaften der Stimme konzentrieren und andere unterdrücken.

2. Begriffsbestimmungen

2.1 Stimme

Die Sprechstimme ist traditioneller Forschungsgegenstand vor allem in Phonetik, Sprechwissenschaft und Sprecherziehung, Rhetorik, Phoniatrie.

Seit einigen Jahren widmet ihr die Geisteswissenschaft verstärkte Aufmerksamkeit, so z.B. aus philosophischer Perspektive (vgl. z.B. Connor 2004; Krämer 2005, 2006 und in diesem Band; Mersch 2006; Waldenfels 2006), aus kultur- und theaterwissenschaftlicher Perspektive (vgl. etwa Chion 2003; Kolesch 2005, 2006a, b; Meyer-Kalkus 2001, 2004), aus medien- und kommunikationswissenschaftlicher Perspektive (vgl. beispielsweise Hartel/Kaspar 2004; Peters 2004; Westphal 2002), aus linguistischer Perspektive (vgl. z.B. Bogner 2000 und 2005; Dolar 2004; Drumbl 2005; König/Brandt 2006) und aus anthropologisch-evolutionärer Perspektive (vgl. z.B. Fitch 2004).

Viele dieser Autoren weisen zu Recht darauf hin, dass die Stimme vielschichtig und ambivalent ist, ein sog. ‚Schwellenphänomen‘: „Denn die Stimme ist immer zweierlei: sie ist sinnlich und sinnhaft, somatisch und semantisch, diskursiv und ikonisch, indexikalisch und symbolisch, natürlich und künstlich, affektiv und kognitiv, physisch und psychisch [...] Körper und Geist, Natur und Kultur, Gefühl und Intellekt“ (Krämer 2005, S. 88); sie ist „individuell und sozial, [...] Aktion der Sprecherin und Passion des Ohres“ (Krämer 2006, S. 12). Nach Dolar (2004, S. 215) ist „die Stimme als Stütze des Sprechens [...] allgegenwärtig und trivial“, gelegentlich wird sie als ‚Nukleus der Geistes-, Human- und Kunstwissenschaften‘ bezeichnet (Kolesch/Krämer 2006, S. 7) oder sogar als Nukleus aller Wissenschaft (Peters 2004, S. 85) und tendenziell als Metapher für Sprechstile (im Sinne Bachtins, vgl. z.B. 1972) oder sogar für Sprechen an sich verwendet. Damit sind Schwierigkeiten bei der begrifflichen Bestimmung und Abgrenzung, vor allem aber für die empirische Beschreibung verbunden.

Die Sprechwissenschaft operiert dagegen mit einem vergleichsweise engen, anatomisch-physiologischen Begriff von Stimme, im Sinne einer Körperfunktion, als Muskelaktivitätsmuster unterschiedlicher Ausprägung, Gestalt und Funktion:

- In engem Bezug zur klinischen Sprechstimmendiagnostik und Phoniatrie wird Stimme betrachtet als Organ; Gegenstand sind z.B. die Physiologie und Anatomie des Kehlkopfes, des Ansatzrohres, der Resonanzräume. Untersucht wird die gesunde und kranke Stimme.
- In engem Bezug zur Phonetik werden Stimmgebung und -wirkung, stimmlich-artikulatorische Ausdrucksformen merkmalsanalytisch auditiv und akustisch beschrieben.
- In engem Bezug zur Linguistik, Rhetorik und Soziophonetik wird Stimme betrachtet als Resultat der Stimmgebung; Gegenstand sind z.B. die Sprechwirkung und die interaktive Ausarbeitung der Stimme in ihrer Verworfenheit mit sprachlichen und körperlichen Ausdrucksformen. Hierbei wird die Stimme als Bestandteil des ‚stimmlich-artikulatorischen Ausdrucks‘ aufgefasst und untersucht, insofern als Teil von persönlicher und sozialer Identität, als Trägerin von ästhetisch-künstlerischem und emotionalem Ausdruck.

- Einen praktischen bzw. praxisbezogenen Zugang zur Stimme hat die Sprecherziehung als angewandte Sprechwissenschaft: In engem Bezug zur Sprechbildung und Gesangspädagogik wird Stimme betrachtet als Tätigkeit der Stimmgebung; Gegenstand sind z.B. die Trainierbarkeit und Optimierung der Laut- und Stimmbildung und der Bezug von Stimme und Körper als unterschiedliche menschliche Ausdrucksdimensionen. Zugrunde liegt zumeist ein physiologisches und ästhetisches Stimmideal westlicher Prägung, das am italienischen Operngesang entwickelt wurde: die anstrengungslose, klangvolle, klare, leistungsfähige Stimme.

2.2 Stimmlich-artikulatorischer Ausdruck

Im oben zitierten Kriminalfall konnte zwar keiner der Ohrenzeugen in den befremdlichen Vokalisationen des Orang Utans ein bekanntes Stimmausdrucksmuster identifizieren, dennoch interpretierten die Zeugen ihren Eindruck von Fremdheit in der gehörten Stimme in Bezug auf den konkreten Situationskontext, als Äußerungen in einer ihnen unbekannten Sprache. Rezipienten versuchen, eine konkrete Stimme bzw. Sprechweise in einer konkreten Situation

vor dem Hintergrund der ‚normalen‘ individuum-spezifischen Stimmlautheit, Stimmhöhe, Klangfarbe, Geschwindigkeit oder Artikulationsweise als eine Variante [zu erfassen], die erst durch die Bedingungen oder Erfordernisse der Situation ausgelöst wurde und die dadurch als situationsbezogener Ausdruck verstanden werden kann. (Stock 1987, S. 57f.)

Für situations- und stimmungsadäquat konventionalisierte Gestaltungsweisen im Sprechschall, mit denen Sprecher u.a. Rollen-, Gruppen-, Handlungsmuster realisieren, existiert in der Sprechwissenschaft der Begriff ‚Stimmlich-artikulatorischer Ausdruck‘ oder auch ‚Sprechausdruck‘ (z.B. Stock 1987, Gutenberg 1998, Bose 2003). Er ist weniger biologisch, sondern stärker sozial und kulturell determiniert und eng verwoben mit der sprachlichen Ebene des Sprechens.

Grundlage bildet der sog. physiognomische Stimmausdruck,

der physiologisch bedingte (durch individuelle Besonderheiten im Bau des Kehlkopfes und des Ansatzrohres) bzw. habitualisierte (durch gewohnheitsmäßige Einstellung der Muskelspannung in den Artikulationsorganen und des subglottalen Luftdrucks) quasikonstante Stimmklang mit seiner individuell spezifischen Lautheit und Tonhöhe [...sowie] jene Modifikationen [...], die bei starken Emotionen wie Wut, Ekel, Lust hervorgebracht werden, oft für die Dauer des Kommunikationsereignisses konstant bleiben und mit instinktmäßiger oder reflektorischer Aktivität zu erklären sind. (Stock 1987, S. 75)

Durch Entfaltung und Ergänzung dieser angeborenen Grundmuster werden im Verlaufe der Kommunikationsbiografie stärker intendierte stimmlich-artikulatorische Ausdrucksmuster erworben: differenziertere, strukturiere Formen, die „bei der Äußerungsgenerierung in Einheit mit dem Wortlaut und als notwendiger ‚Kommentar‘ zu diesem geplant und realisiert“ werden (Stock/Suttner 1991, S. 64).

Die Verwendung des Begriffs ‚Ausdruck‘ verweist auf die Tradition der ausdruckspsychologischen Sprechstimmforschung (stellvertretend z.B. Werner 1932; Bühler 1933 und 1934; Trubetzkoy 1939; Trojan 1948; Kirchoff 1965), welche stimmliche Äußerungsformen als Ausdruck, als äußere Erscheinungsformen innerer Veranlagungen, Empfindungen und Leidenschaften deutet. Nach Görlitz (1972, S. 62) deckt der Begriff allgemein „viele von dem, was alltäglich als ‚wie jemand spricht‘ begegnet“. Allerdings ist der Ausdrucksbegriff wegen seiner Polyvalenz bis hin zur Überdehnung und wegen seiner Äquivalenzen oft kritisiert worden (vgl. stellvertretend die Diskussion in Görlitz 1972; Rothe 2006). Stimmlich-artikulatorischer Ausdruck wird einerseits verstanden als die Realisierung von Expressivem (z.B. Trojan 1948), andererseits als Oberbegriff, als die Realisierung von Rollen-, Gruppen-, Handlungsmustern im Sprechschall (z.B. Geißner 1981). Ausdruck ist

im phänographischen Zugang ein unmittelbar Antreffbares, ist unvermitteltes Erscheinen, während Ausdruck im funktional-sematologischen Zugang ein als Zeichen fungierendes Medium ist, das für etwas anderes (Befindlichkeit, Positionalität, personales Sosein) steht. Die konditionale Fassung schließlich bezieht sich auf die Ausdruck konstituierenden Faktoren. (Görlitz 1972, S. 55; ähnlich auch Kirchoff 1991)

Interessant ist der Ausdrucksbegriff, weil er nicht nur auf das Individuum verweist, das seinen Empfindungen willkürlich oder unwillkürlich Ausdruck gibt, sondern auch auf den Adressaten, bei dem Empfindungen erregt werden, und damit auf den Interaktionsprozess; Ausdruck ist immer Ausdruck von jemandem für jemanden in einer konkreten Situation.

Wahrnehmung und Deutung des stimmlich-artikulatorischen Ausdrucks enthalten Urteile sowohl über den physiognomischen als auch über den pathognomischen Stimmausdruck, weil Sprechen beides enthält, nämlich „immer auch Ausdruck habitueller kognitiver, emotionaler und voluntativer Geprägtheiten des Sprechenden [...], immer auch Ausdruck und Vollzug okkasioneller kognitiver, emotionaler und voluntativer Prozesse“ (Gutenberg 2001, S. 73). Die Bewertung des pathognomischen Ausdrucks hängt von der des physiognomischen ab. Über den Vergleich mit Ausdruckserscheinungen anderer Sprecher in ähnlichen Situationen „bilden sich innerhalb eines Kulturkreises, einer Sprachgemeinschaft oder einer sozialen Gruppe bestimmte Erwartungsvorstellungen über das mögliche bzw. zulässige Ausdrucksverhalten heraus“ (Stock 1987, S. 57 f.). Den Ausdrucksmus-

tern entsprechen also gesellschaftlich bedingte Hörmuster, interiorisierte Normen bzw. Erwartungsmuster zur Angemessenheit stimmlich-artikulatorischer Gestaltung als situationsadäquate Realisation bestimmter Sprechhandlungs-, Rollen-, Sozialmuster (vgl. Gutenberg 2001; ähnlich Geißner 1984; Selting 1997). Mit eindeutigen Zuordnungen ist allerdings nicht zu rechnen, denn Musterbildungen der Sprechfähigkeit sind heterogen, können sich also wechselseitig überschneiden und einander interdependent durchdringen. Kinder erwerben Hör- und Ausdrucksmuster gemeinsam mit der Sprache anhand von Wahrnehmung, Abstraktion und kreativer Umsetzung der jeweils individuell ausgeführten Ausdrucksweisen der kommunikativen Umgebung.

Bei der auditiven Wahrnehmung ist nach Bühler (1927) der ganze psychophysische Organismus beteiligt, und Hörer vergegenwärtigen sich beim Wahrnehmen von Stimmen die ganze Person. Was als Ausdruck hervorgebracht wird, löst einen Verhaltensimpuls im Adressaten aus, kann in seiner Bedeutung nicht ohne dessen Reaktionen verstanden werden, vgl. z.B. Trubetzkoy (1939 [1989], S. 18):

Wenn wir jemanden reden hören, so hören wir, wer spricht, in welchem Tone er spricht und was er sagt. Es liegt ja eigentlich nur ein einziger akustischer Eindruck vor. Wir zerlegen ihn aber in seine Bestandteile, und zwar immer vom Standpunkte der drei Bühlerschen Sprachfunktionen: gewisse Eigenschaften des wahrgenommenen Schalles fassen wir auf als Kundgabe, als Symptom des Sprechers (z.B. seine Stimmlage), gewisse andere als Mittel zur Auslösung bestimmter Gefühle beim Hörer, und endlich noch andere als Merkmale, an denen die Wörter mit bestimmten Bedeutungen und die aus ihnen bestehenden Sätze erkannt werden.

In ihrem Eindruck reproduzieren Rezipienten nicht einfach das vom Sprecher Ausgedrückte, sondern ziehen Rückschlüsse auf zugrundeliegende Befindlichkeiten, und zwar vor einem situativen, historischen und kulturellen Horizont von Traditionen, Praktiken, Medien, Kultur- und Kunstformen. Eindruck umfasst demnach sowohl die Wahrnehmung von Ausdruck als auch die Verknüpfung des Wahrgenommenen mit eigenen Assoziationen und die Musterzuschreibung durch den Vergleich mit Erinnerungen – also die Ergänzung um Informationen, die nicht unmittelbar im Geäußerten enthalten sind (Bühler 1933). Rezipienten identifizieren Ganzheiten, die im Verlauf der Kommunikation als Sukzessivgestalten erscheinen. Dadurch weist die Wahrnehmung bzw. Deutung stimmlich-artikulatorischen Ausdrucks individuelle Spielräume, Varianten, Unschärfen bis hin zu Irrtümern auf, dennoch ist sie ein unverzichtbares Mittel der sozialen Orientierung.

2.3 Parameter des stimmlich-artikulatorischen Ausdrucks

Stimmlich-artikulatorischer Ausdruck stellt einen Merkmalskomplex aus stimmlich-melodischen, aber auch aus temporalen, dynamischen und arti-

kulatorischen Parametern dar, die an auditiv wahrnehmbaren physiologisch-akustischen Grundlagen der Stimmproduktion und Artikulation orientiert sind. Allerdings ist der Zusammenhang zwischen akustischen Grundlagen und Hörempfindungen nicht linear. Bezugsgrößen sind deshalb keine (absoluten, objektiven) Messwerte, sondern konsequent auditive Eindrücke. Entsprechende Termini und Beschreibungssysteme stammen aus der ausdruckspsychologischen Sprechstimmforschung (z.B. Trojan 1948 und 1975; Fährmann 1960; Höffe 1965), der Sprechwissenschaft (z.B. Winkler 1969; Krech/Richter/Stock/Suttner 1991; Stock 1996; Gutenberg 1998; Heilmann 2001), der klinischen Sprechstimmforschung (z.B. Pfaü/Streubel 1982; Seidner/Wendler 1982; Vieregge 1989 und 1996; Wendler/Seidner/Kittel/Eysoldt 1996) und aus der Phonetik (z.B. Laver 1980 und 2000; Terhardt 1998; Neppert 1999; Kent/Ball 2000; Ketzmerick 2007).

Stimmlich-artikulatorischer Ausdruck ist beschreibbar anhand folgender Parameter (ausführlich hierzu Bose 2001 und 2003 und die dort genannte Literatur):

- *Sprechtonhöhe* (hoch *vs.* tief, bewegt *vs.* monoton), abhängig von Periodizität und der Periodendauer der Schallsignale, aber auch von Spektralmerkmalen der Stimme. Physiologisch entstehen Tonhöhenveränderungen durch Veränderungen in der Stimmlippenspannung, verschiedene Einstellungen des Spannapparates der Kehlkopfmuskulatur und Veränderungen des subglottischen Drucks.
- *Stimmklang*, abhängig von der Tätigkeit der Glottis, der Klangbildung und Resonanzwirkung im Ansatzrohr und der Atmung. Zum Stimmklang gehören Merkmale wie
 - *Klangfülle* (klangvoll *vs.* klangarm), physiologisch bedingt durch die Regelmäßigkeit im Schwingungsablauf der Stimmlippen und durch Resonanzeigenschaften des Ansatzrohres, die die Teiltonstruktur des Spektrums bestimmen.
 - *Klangfarbe* (hell *vs.* dunkel), physiologisch bedingt durch die Formung des Ansatzrohres, aber auch durch die Tonhöhe.
 - *Faukale Distanz* (weit *vs.* eng), physiologisch bedingt durch die Gestaltung des Rachenraumes, genauer durch die Weite bzw. Enge im Bereich der hinteren Gaumenbögen, durch die Lage der Zungenwurzel und die Lage des Kehlkopfes während der Stimmgebung,
 - *Stimmein- und -ausschwingphasen* (weich *vs.* hart), physiologisch bedingt durch die Stimmlippenspannung und die Stärke des subglottischen Drucks.
 - *Geräuschanteil* (klar *vs.* geräuschhaft, also z.B. behaucht, rau, gepresst, geknarrt), physiologisch bedingt durch die Regelmäßigkeit bzw. Unregelmäßigkeit der Stimmlippenschwingungen.

- *Lautheit* (laut *vs.* leise), abhängig von der physikalisch messbaren Schall-Amplitude. Physiologisch entstehen Lautheitsveränderungen durch Veränderungen des subglottischen Drucks, aber auch durch Veränderungen im Schwingungsablauf der Stimmlippen.
- *Sprechgeschwindigkeit* (schnell *vs.* langsam), verstanden als Realisierung bestimmter sprachlicher Einheiten (Laute/Silben/Wörter) in einer bestimmten Zeiteinheit (s/min). Ihre auditive Wahrnehmung bzw. Beurteilung hängt aber auch von Sprechpausen, Akzentuierungsstärke, Artikulationspräzision, Lautdehnungen usw. ab.
- *Akzentuierung* (stark *vs.* schwach, eher melodisch *vs.* dynamisch *vs.* temporal), beschrieben anhand der Position, Stärke und Form auditiv wahrnehmbarer Akzentstellen.
- *Sprechrhythmus* (staccato *vs.* legato, skandierend *vs.* nicht skandierend), verstanden als zeitgliedernde Erscheinung, die die Regelung des Wechsels zwischen akzentuierten, nicht akzentuierten und schwach akzentuierten Silben betrifft und sich mit den Kategorien ‚Spannung‘ und ‚Entspannung‘ beschreiben lässt.
- *Sprechspannung* (gespannt *vs.* ungespannt), abhängig von Stimmlippen-spannung, Expirationsdruck und relativer Spannung der Artikulationsmuskulatur.
- *Artikulation*, verstanden als Realisierung der Sprechlaute. Bezugsgrößen sind Aussprachenormen ebenso wie emotionale Gestimmtheit. Zur Artikulation gehören Merkmale wie:
 - *Artikulationspräzision* (präzise *vs.* unpräzise), der Sorgfältigkeitsgrad bei der Ausspracherealisation, abhängig von Artikulationsspannung, Lippentätigkeit, Kieferöffnungsweite, Zungenlage, Gaumensegelfunktion, Kehlkopfstand.
 - *Lippenstellung* (gestülpt *vs.* gespreizt), die auditiv auffällige Überformung der Artikulation durch Lippenbewegungen, die nicht ursprünglich zur Schall- bzw. Lauterzeugung gehören.
 - *Lautdauer* (gedehnt *vs.* verkürzt) in Bezug auf die Position innerhalb der Sprachäußerung, abhängig von der Sprechgeschwindigkeit.

Wie bereits aus dieser knappen Darstellung hervorgeht, kommen Beschreibung und Interpretation der einzelnen Parameter aufgrund komplexer, nicht immer genau aufzuschlüsselnder Höreindrücke zustande, sind also nicht eindimensional (die Einschätzung von Stimmklangfarbe ist z.B. nicht unabhängig von der Sprechstimmhöhe; die Einschätzung der Sprechstimmlagen geschieht auch über das Nachempfinden von Spannungsverhältnissen usw.). Ebenso ist nicht klar zu trennen zwischen i.e.S. ausschließlich stimmlichen *vs.* prosodischen *vs.* artikulatorischen Gestaltungsweisen.

3. Empirische Erforschung des stimmlich-artikulatorischen Ausdrucks

Nicht nur theoretisch, sondern auch methodisch stehen viele neuere empirische Arbeiten zur Beschreibung, Klassifikation und Wirkung stimmlich-artikulatorischen Ausdrucks in der Tradition der ausdruckspsychologischen Sprechstimmforschung. Deshalb werden zunächst stellvertretend zwei einflussreiche ausdruckspsychologische Arbeiten vorgestellt, danach exemplarisch Vorgehensweisen und Erkenntnisse der Sprechwissenschaft bei der empirischen Erforschung des sprechstimmlichen Ausdrucks. Ausgewählt wurden sowohl Arbeiten, in denen Merkmale beschrieben und klassifiziert werden (Stimmphysiologie-, Emotionsforschung), als auch Arbeiten, in denen der stimmlich-artikulatorische Ausdruck in seiner Wirkung auf Hörer (Sprechwirkungsforschung) und in seiner interaktiven Ausarbeitung (Gesprächsforschung) betrachtet wird.

3.1 Ausdruckspsychologische Sprechstimmforschung

Ziel ausdruckspsychologischer Sprechstimmforschungen ist die Beschreibung des stimmlich-artikulatorischen Ausdrucks anhand von phonischen Merkmalen, die Klärung seiner Funktion in der mündlichen Kommunikation und die Bestimmung der Bedingungen für bestimmte Gestaltungsmuster. Gemeinsam ist ihnen die Verwendung experimentell erstellten bzw. konstruierten Materials in kontrollierten Situationen. Aus der ausdruckspsychologischen Sprechstimmforschung sind Konzeptualisierungen zum stimmlich-artikulatorischen Ausdruck und typologisch ausgerichtete Ordnungsversuche hervorgegangen, darüber hinaus Beiträge zur Kultivierung und ästhetischen Normierung des sprechstimmlichen Erscheinungsgeltes (vgl. hierzu die Diskussion in Görlitz 1972; Meyer-Kalkus 2001; Rothe 2006). Vorgestellt werden eine Perzeptionsstudie (Bühler 1933) und eine merkmalsanalytische Arbeit (Trojan 1948).

Zusammen mit Kollegen und Schülern hat Bühler eine bemerkenswerte Massenbefragung zum „Zusammenhang von Stimme und Persönlichkeit“ durchgeführt (Bühler 1933; Herzog 1933), in der statt des klassischen Laborversuchs das neue Medium Rundfunk genutzt wurde. Im sog. physiognomischen Radio-Experiment sollte geklärt werden, welche Informationen Hörer aus einer Radiostimme entnehmen können und inwiefern sie in ihren Urteilen übereinstimmen. Neun Sprecher/innen unterschiedlichen Alters aus allen Schichten der Wiener Bevölkerung wurden gebeten, eine Ballade und einen Sachtext zu sprechen. Diese Sprechfassungen wurden zusammen mit einem Fragenkatalog im Radio gesendet. 2700 Radiohörer gaben ihren Eindruck darüber kund, ob ihnen die jeweilige Stimme sympathisch war, welche Vermutungen sie über Geschlecht, Alter, Aussehen,

Beruf der Sprecher/innen hatten u.a.m. Die Auswertung ergab eine weithin eindeutige Identifizierung des Geschlechts der Sprechenden, eine große Übereinstimmung in der Einschätzung von Größe und Gestalt, dagegen Unsicherheiten in der Bestimmung von Alter und Beruf.

Bühler (1933 [1968], S. 192 ff.) schlussfolgert, dass Rezipienten mit einem Mechanismus der Unterstellung arbeiten, d.h. dass sie ihre Interpretationen über einen unbekanntem Sprecher aus dem Vergleich mit Eigenschaften einer bekannten Person gewinnen, deren Stimme Ähnlichkeiten mit der gehörten aufweist. Nach Bühler funktioniert dieser Mechanismus der Unterstellung als Resonanzverfahren (Einfühlung bzw. unwillkürliche Ansprechbarkeit durch den Ausdruck anderer Menschen) und als Indizienverfahren (Spurenlesen, um das auf indirekte Weise Ausgedrückte zu verstehen). Damit misst Bühler der Subjektleistung bei der Konstitution von Wahrnehmungseindrücken eine entscheidende Rolle zu. Die Bedeutungszuschreibung in der Kommunikation fasst er als Top-Down-Verarbeitung unter Mitwirkung zentraler Antizipationen auf und betont die Relevanz von Konventionen und von Situationsindizien (vgl. hierzu auch die Diskussion in Ehlich 2004; Nothdurft 2004; Hirschfeld/Stock ersch. 2009). Kommunikationsbeteiligte ziehen demnach ihre Schlüsse darüber, wie etwas, was der andere sagt bzw. meint, zu verstehen und zu interpretieren ist, aufgrund von Hinweisreizen, Kontexthilfen und Situationsindizien, zu denen auch der stimmlich-artikulatorische Ausdruck gehört. Allerdings sind diese Hinweisreize nicht zweifelsfrei zu identifizieren oder vorherzusagen, weil sich die Kommunizierenden aus dem Gesamtkomplex der Kommunikation bedienen und das nicht immer auf die gleiche Weise (Bühler 1933; vgl. hierzu auch Stock 1991; Hirschfeld/Stock ersch. 2009).

Trojan (1948), Bühler-Schüler und Sprecherzieher, hat mit Rückgriff auf die Phonologie Trubetzkoy's (1939) eine Typologie des sprechstimmlichen Ausdrucks vorgelegt, die sog. ‚phonetische Lautstilistik‘, verstanden als eine Art zweites semantisches Zeichensystem. Gegenstand der Untersuchung sind Gestaltungs- und Formungsregeln des sozial adressierten Ausdrucks und das Verhältnis von Ausdrucksmitteln und -inhalten. Trojan nimmt sog. Schallbilder der gesprochenen Sprache an, Zeichen mit Ausdrucks- und Appellfunktion: Schallbilder

stellen meist unselbständig auftretende, phonisch-artikulatorische Zeichen dar, die in der Regel instinktiv gesendet werden, koexistierend angeordnet sind und in der Hauptsache existentielle Vorgänge im Zeichensender ausdrücken oder vorgeben; sie haben sich aus vorsprachlichen Rudimentärformen entwickelt, sind mit der Sprache eine sehr innige Fusion eingegangen und haben sich eben dadurch außerordentlich verfeinert und differenziert. (Trojan 1948, S. 45)

Die Schallbilder sind nicht ausschließlich reflektorisch oder affektiv bedingt, sondern können auch intendiert und situationsbezogen verwendet (und unterdrückt) werden, und sie unterliegen in gewissem Umfang den glei-

chen Bildungsprinzipien wie die sprachlichen Zeichen (Trojan 1948., S. 2). Den Schallbildern liegen interne Normen zugrunde, die sog. Akueme, wobei die konkreten Realisationen „innerhalb eines Spielraumes, den ihnen eben diese Normen gewähren, eine gewisse Freiheit und Selbständigkeit [bewahren]“ (ebd., S. 11). Um die Akueme zu beschreiben, sucht Trojan aus pathognomischer Perspektive nach Konstanz des Affektausdrucks bei verschiedenen Individuen, aus physiognomischer Perspektive sucht er beim Ausdruck unterschiedlicher Affekte nach personeller Konstanz, die durch „Bau und Zustand des Körpers, Bevorzugung einer bestimmten seelischen Grundhaltung“ bedingt ist (ebd., S. 15).

Material für die Untersuchung bildete ein Satz (*Und das ist von nun an mein Tun und Treiben*), der von drei Schauspielern in zehn vorher festgelegten Schallbildern (z.B. ‚Angst, Ekel, freudige Rührung, Langeweile, zielbewusster Wille‘) gesprochen wurde. Die sprechstimmlichen Realisationen wurden vom Autor auf ihre Treffsicherheit beurteilt und von geschulten Hörern auditiv-akustisch beschrieben: mittels eines Beurteilungskatalogs mit phonetisch-physiologisch basierten Merkmalen, wie z.B. Atmungsform, Stimmregister, Akzentform, Artikulationsspannung, Stimmeinsatz u.a.m. (Trojan 1948, S. 104 ff.). Ergebnisse dieser Untersuchungen sind eine Merkmalsammlung zur phonetisch-physiologisch basierten Beschreibung der Schallbilder und Lehrsätze einer Stildiagnostik auf sprachinhaltlicher und sprechexpressiver Grundlage.

Durch ein solches Vorgehen können zwar wichtige Erkenntnisse über phonische Auslösefaktoren und dadurch erzeugte Wirkungen gewonnen werden, aber zur interaktiven Relevanz des stimmlich-artikulatorischen Ausdrucks im Kommunikationsprozess sind nur eingeschränkte Aussagen möglich. Jüngere Untersuchungen über Form-Funktions-Beziehungen im stimmlich-artikulatorischen Ausdruck, z.B. zur Sprechwirkung, sind zwar ebenfalls als Kommunikationsexperimente mit konstruiertem Material angelegt, aber mit theoretisch veränderten Konzepten und verfeinerten Methoden.

3.2 Stimmphysiologie und -pathologie

Phoniatrie, Logopädie und klinische Sprechwissenschaft befassen sich mit der Diagnostik und Therapie der Sprech- und Singstimme. Voraussetzung ist eine möglichst klare und transparente Beschreibung stimmlicher Eigenschaften, sowohl apparativ als auch auditiv. Da für die meisten akustischen Erhebungen gehaltene Vokale mit ausreichender Dauer produziert werden müssen, sind diese Verfahren für die Beschreibung stimmlicher Qualitäten bei fortlaufendem Sprechen (noch) nicht geeignet (vgl. stellvertretend Nawka/Anders 1996; Anders 1997). Deswegen spielt in der klinischen Praxis die auditive Beurteilung nach wie vor eine sehr große Rolle. Sie bietet die ein-

fachste und schnellste Möglichkeit, eine Stimmstörung verlässlich zu bewerten bzw. zu dokumentieren (vgl. stellvertretend Schneider/Bigenzahn 2007).

Allerdings steht für die Beschreibung von Hörwahrnehmungen der Stimme nur ein sehr begrenzter Wortschatz zur Verfügung:

Wenn man sich exklusiv auf auditive Kategorien beschränkt, bleibt letztlich ein Adjektivpaar übrig, nämlich laut und leise. Alle anderen Kennzeichnungen, wie hoch oder tief, hell oder dunkel, strahlend oder stumpf, sind metaphorische Umschreibungen mit Anleihen aus anderen Sinnesbereichen. Der Mangel an spezifischen sprachlichen Ausdrucksmöglichkeiten für gehörte Eindrücke führt zu einer großen Vielfalt bei der Wahl geeignet erscheinender Wörter. Im Falle des menschlichen Stimmklangs kommt erschwerend hinzu, daß dieser selbst in außergewöhnlicher Vielfalt auftritt, mit individuellen Ausprägungen, wie sie von Gesichtern und Fingerabdrücken bekannt sind. (Wendler 1996, S. 3)

Aus der therapeutischen Erfahrung heraus sind zahlreiche Kataloge zur Beurteilung von stimmlichen Eigenschaften und Leistungen entstanden, die oft eine Vielzahl von solchen metaphorischen Umschreibungen enthalten (vgl. z.B. angenehm, schrill, metallisch, spitz, hoch/tief; klein/groß, warm/kalt, schneidend, rund, glatt). In einem Überblick über solche Systeme nennt Sonninen (1970) z.B. allein zur Beschreibung heiserer Stimmen 59 Adjektive. Aufgrund ihrer Idiosynkrasie und Unübersichtlichkeit haben solche Kataloge nur eingeschränkte Verbreitung gefunden.

3.2.1 Parameterbasierte Beschreibungssysteme

Für die klinische Praxis wurden deshalb transparente Beurteilungssysteme für stimmliche Eigenschaften entwickelt, die sich auf Erkenntnisse über die Anatomie, Physiologie und Pathologie der Stimme stützen. Stellvertretend wird auf drei international anerkannte und verbreitete Systeme zur Beurteilung einer Stimmstörung verwiesen. Gemeinsam ist ihnen die Orientierung an klar definierten Bewertungssystemen und die Konzentration auf wenige wesentliche Klangphänomene der Stimme, die skaliert beschrieben werden. Anhand dieser Systeme sind zunächst grobe Beschreibungen und Kategorisierungen von Stimmen möglich, an die sich exakte Diagnosen anschließen können.

Die sog. *GRBAS-Skala* (Hirano 1981) stützt sich auf fünf Parameter: G – *overall grade for degree* (Gesamtgrad der Störung, Instabilität), R – *rough* (Rauigkeit), B – *breathy* (Behauchtheit), A – *asthenic* (Asthenie, Schwachheit), S – *strained quality* (Spannung, Gepresstheit). Der Ausprägungsgrad dieser Parameter wird auditiv mit Hilfe einer vierstufigen Skala bestimmt: 0 – nicht vorhanden, 1 – geringgradig, 2 – mittelgradig, 3 – hochgradig vorhanden.

Da sich in der Anwendung des Systems zwei Parameter als uneindeutig erwiesen (Parameter A wegen seiner Korrelation vor allem mit B und Parameter S wegen zu stark streuender Beurteilungen), wurde diese Skala über-

arbeitet und auf drei Faktoren eingeschränkt, z.B. in der sog. *RBH-Klassifikation* (Wendler/Rauhut/Krüger 1986; Nawka/Anders 1996; Nawka/Evans 2005). Sie berücksichtigt ebenfalls vier Ausprägungsgrade und beruht auf den Parametern H – Heiserkeit, R – Rauigkeit, B – Behauchtheit. Maßstab ist die sog. euphone Stimme (R0 – B0 – H0). Heiserkeit wird als „unspezifisches Leitsymptom einer Stimmstörung“ definiert. Sie entsteht durch die beiden Hauptkomponenten „phonatorisch unkontrollierter Luftverbrauch“ (Behauchtheit) und „irreguläre Schwingungen“ (Rauigkeit). Die Heiserkeit wird immer von einer der beiden untergeordneten Dimensionen dominiert, stellt also kein völlig unabhängiges Urteil dar (Nawka/Anders 1996, S. 114).

Ähnlich aufgebaut ist auch das System *Consensus Auditory-Perceptual Evaluation of Voice* (CAPE-V) der *American Speech-Language-Hearing Association* (ASHA) (2002). Es verwendet sechs Parameter: allgemeine Einschätzung der Störung (*overall severity*), Rauigkeit (*roughness*), Behauchtheit (*breathiness*), Gepresstheit (*strain*), Grundfrequenz (*pitch*), Stimmstärke (*loudness*). Der Stärkegrad ist anders als in den anderen beiden Systemen mittels Analogskalen stufenlos anzugeben, darüber hinaus wird vermerkt, ob das betreffende Merkmal konsistent (*consistent*) oder nur gelegentlich (*intermittent*) auftritt.

Heiserkeit ist aufgrund physikalisch messbarer und physiologisch definierbarer Grundlagen des Stimmschalls nicht eindeutig zu beschreiben, da ihr unterschiedliche pathophysiologische Mechanismen zugrunde liegen. Trotz ihrer unterschiedlichen Qualitäten ist sie aber auditiv durchaus erkennbar und klassifizierbar (Evans/Nawka 2007, S. 313). Die auditive Beurteilung von Heiserkeit ist vor allem dann zuverlässig, wenn sie auf Gruppenurteilen nach Hörtraining basiert. Deswegen sehen die drei genannten Systeme auch Hörtrainings vor. Dem RBH-System liegt z.B. eine Übungs-CD mit Klangbeispielen bei, die von Expertengruppen übereinstimmend eingeschätzt worden sind. In der therapeutischen Praxis werden mit der ordinalen Merkmalskalierung sehr einhellige Hörerurteile erzielt, die auch nach langen Zeiträumen reproduzierbar sind.

Auch gesunde Stimmen entsprechen nicht immer den Kriterien einer euphonen Stimme, sondern weisen oft gering- und mittelgradige Heiserkeiten auf, die sowohl habituellen Ursprungs als auch Ausdruck situationsbezogener Stimmverwendung sein können. Deshalb sind solche parameterbasierten Systeme zur Beschreibung und Beurteilung von Stimmklängen durchaus auch außerhalb der therapeutischen Praxis nutzbar, z.B. für die Analyse gesprochener Äußerungen in der Gesprächskommunikation (vgl. ausführlich Bose/Schwarze/Wendt ersch. 2009).

Notwendige Voraussetzung für die Beschreibung und Beurteilung stimmlich-artikulatorischer Ausdrucksformen ist das sog. funktionell-analytische Hören, „die Fertigkeit, Sprechereignisse mit Hilfe des Gehörs segmental und suprasegmental mit Hilfe von Beurteilung, Skalierung und Transkription in Merkmale und Merkmalskomplexe zu zerlegen“ (Vieregge

1996, S. 1). Aufgrund eines intensiven Hörtrainings sind Hörer in der Lage, auf einzelne stimmlich-artikulatorische Merkmale des auditiv wahrnehmbaren Komplexsignals zu achten und sie zu kennzeichnen (Bottom-Up-Hören im Gegensatz zum normalerweise ganzheitlichen, vorrangig semantisch orientierten Top-Down-Wahrnehmungsprozess in der Kommunikation). Mit Hilfe des analytischen Hörens wird das Gehörte auf eine internalisierte Erwartung hin geprüft. Zwar können sich auch trainierte Hörer nie ganz von ihrer Kulturzugehörigkeit, von ihrem Vorwissen über die gehörte Sprache und von ihren Angemessenheitsvorstellungen über stimmlich-artikulatorische Äußerungen lösen, sie haben aber durch Hörtraining gelernt, sich möglichst wenig um diesen ‚semantischen Rest‘ (Vieregge 1996, S. 21) zu kümmern. Analytisches Hören kann immer nur einen relativen, nie einen absoluten Status beanspruchen, deshalb ist eine Arbeit in Hörergruppen sinnvoll. Bei Verwendung eines transparenten Bezugssystems steigt die Zuverlässigkeit auditiver Beschreibungen mit der Übereinstimmung unter den Gruppenmitgliedern. Das Maximum an Objektivität einer auditiven Beschreibung stellt die sog. Konsensusbeschreibung dar, mit der alle Gruppenmitglieder einverstanden und von deren Objektivität sie überzeugt sind (vgl. ausführlich Bose/Schwarze/Wendt ersch. 2009). Auch bei einem solchen Verfahren ist natürlich nicht auszuschließen, dass verschiedene Gruppen zu verschiedenen Ergebnissen kommen können.

3.2.2 Kulturgebundenheit von Stimmbewertungen

Auf die Kulturgebundenheit von stimmlich-artikulatorischem Ausdruck und seiner Beurteilung weisen Ergebnisse aus Studien zum Stimmgebrauch in anderen Kulturen hin. So hat Grawunder (2009) in einer Feldstudie die Anatomie, Physiologie und Akustik des südsibirischen Kehlgesangs untersucht, einer kulturgebundenen kunstvollen Stimmproduktion. Vergleichbar dem westlichem Obertongesang werden einzelne Obertöne verstärkt und bilden die Melodie. Wesentlich für den Kehlgesang ist aber das besondere Timbre, das sich aus dem Zusammenspiel von verstärkten Obertönen und zwei grundlegenden Phonationsmodi mit starker laryngaler Verengung ergibt: einer gespannten, fast gepressten Stimmqualität und einer tief rau pulsierenden Stimmqualität. Aufgrund des eigenen physiologischen Optimums des Kehlgesangs ist seine Untersuchung grundsätzlich interessant für das Verständnis von Stimmproduktion an sich.

25 Kehlsänger wurden vor Ort mit einer nichtinvasiven Methode untersucht, der Endoskopie (mittels Drei-Kanal-Aufnahmen: Mikrophon-, Elektrolottografie-, Brustresonanzsignal). Nach Grawunder (ebd., S. 35) sind die phonetischen Termini, die im westlichen Kulturkreis zur Beschreibung für solche Stimmqualitäten zur Verfügung stehen (gepresst, rau), eng verknüpft mit Beschreibungen von Stimmstörungen. Insofern weicht der

Kehlgesang sehr stark vom westlichen Stimmideal ab und könnte potentiell stimmgefährdend sein. Die Untersuchung zeigt allerdings, dass gute Sänger den Kehlgesang jahrzehntelang professionell produzieren können, ohne Stimm Schäden davonzutragen (was selbstverständlich ebenfalls für gute Sänger im westlichen Kulturraum gilt, die extreme Stimmklänge nutzen, wie z.B. im Heavy Metal). Darüber hinaus weist Grawunder (2009., S. 79) darauf hin, dass die im Kehlgesang verwendeten Phonationsarten in einigen Sprachen auch linguistisch mit distinktiver Funktion genutzt werden, z.B. im rauen, gepressten Tieftonregister des Bai (Sinutibetisch) (Esling 2002) oder im Phonationstyp ‚strident‘ des !Xoon (Khoisan) (Traill 1986).

Gauß (2008 und 2009) hat den Bereich der praktischen Sprechbildung in verschiedenen Kulturen betrachtet. Unter Sprechbildung wird die Ausbildung von Körper, Stimme, Haltung und Artikulation in Bezug auf konkretes situationsadäquates und partnerbezogenes Sprechhandeln verstanden, unter Stimmbildung speziell die Ausbildung einer tragfähigen Stimme und deutlichen Artikulation für sprechintensive Berufe und sprechkünstlerische Tätigkeit. Ziel dieser ‚westlichen‘ Stimmbildung (McAllister-Viel 2006) ist es, die Stimme leicht und entspannt (‚frei und natürlich‘) zu produzieren, ohne jede übermäßige Anstrengung, die Stimmbelastungen und -schäden oder zumindest Klangverzerrungen zur Folge hat. Basis dieses Stimmideals wie auch seiner didaktischen Vermittlung ist die Orientierung an Anatomie und Physiologie des Sprechens. Methodisch bilden Körper- und Stimmtraining eine Einheit, z.B. durch Balance-, Resonanz-, Impuls-, Spannungs-/Entspannungsübungen. Partner- und Situationsbezug des Sprechens, die sog. Ansprechhaltung, werden vermittelt nach dem Prinzip des ‚gestischen Sprechens‘, z.B. durch die Entwicklung möglichst konkreter, klarer Vorstellungen und Aktivierung entsprechender Erlebnisinhalte. Ausdruck und innere Haltung werden also als Einheit begriffen (vgl. Klawitter/Minnich/Honigmann/Mont'agh 1981; Aderhold 1993; Klawitter/Köhler 2004).

Eine solche Vorstellung von einer guten Stimme gilt durchaus nicht für alle Kulturen, wie Gauß das für die traditionelle koreanische Ausbildung des sog. P'ansori zeigt. P'ansori ist eine Erzähl- und Gesangskunst, deren stimmliche Interpretation bis in detaillierte Klangfarben und melodische Ornamentierungen hinein festgelegt ist und die allein über auditive Nachahmung gelernt wird. Aus einer vergleichenden Analyse beider sprechbildnerischer Konzepte (McAllister-Viel 2006) geht hervor, dass damit jeweils unterschiedliche psychophysische Konzepte verbunden sind, d.h. Vorstellungen davon, was der Mensch ist, wie sein Körper und seine Stimme funktionieren bzw. wie sie zu gebrauchen sind und mit welchen Methoden sie zu formen sind. Während die Grundlage der westlichen Sprechbildung in der schulmedizinischen Auffassung von Stimmhygiene zu suchen ist, liegt sie für das P'ansori in der Körperauffassung der chinesischen Medizin.

Trainingstradition und Ästhetik des P'ansori rütteln nach Gauß (2009, S. 44) am

Dogma der westlichen Sprechbildung und ihren medizinischen Kategorien. Zwar hat das intensive Training auf den Stimmbändern der P'ansorisängerin physiologische Spuren hinterlassen und die Alltagsstimme einen heiseren Klang, dennoch kann nicht von ‚Stimmschädigung‘ gesprochen werden, wenn es sich im Sinne der Kultur um eine ‚ausgebildete‘ Stimme handelt, der Klang der Stimme geschätzt wird und die Alltagskommunikation nicht eingeschränkt ist.

Die Autorin (ebd.) macht außerdem darauf aufmerksam, dass sich auch innerhalb der westlichen Tradition verstärkt Stimmkultivierungsmethoden etablieren, die übermäßige Spannung und extreme Stimmlängen nutzen (vgl. hierzu z.B. auch Meyer-Kalkus 2004; Kolesch 2005).

Wie ein erweiterter ethnografischer Blick zeigt, sind Anatomie und Physiologie als grundlegende Bezugsgrößen der Sprechbildung nicht universell. Trotz gemeinsamer biologischer Basis der Stimmfunktion entwickeln sich in verschiedenen Kulturen Stimmgebrauch und -beurteilung nicht unabhängig von kulturell-historisch geprägten Auffassungen zu Körperlichkeit und Körpergebrauch (vgl. dazu u.a. auch Bogner 2000 und 2005; Waldenfels 2006; Mersch 2006). Für Bogner (2005, S. 109) ist die Kulturgebundenheit der Stimme deshalb auch Gegenstand einer interkulturellen Linguistik. Wie die Arbeiten von Grawunder (2009) und Gauß (2009) zeigen, sind kulturvergleichende oder interkulturelle Forschungen zu diesem Bereich als komplexer hermeneutischer Verstehensprozess mit verschiedenen Methoden zu konzipieren (ethnografisch basierte Feldforschung, Teilnahme an sprechbildnerischer und sprechkünstlerischer Ausbildung, Expertengespräch, terminologische und phonetische Analyse). Nur so sind auch die eigenen kulturellen Prägungen der wissenschaftlichen Beschreibung von stimmlich-artikulatorischem Ausdruck aufzudecken. Solchermaßen gewonnene Erkenntnisse haben auch sehr konkrete Auswirkungen, z.B. auf die Analyse und Notation stimmlich-artikulatorischen Ausdrucks in Gesprächsforschungen, welche Parameter überhaupt als relevant wahrgenommen werden und wie sie zu notieren sind.

3.3 Stimmlich-artikulatorischer Emotionsausdruck

Empirische Studien zum sprechstimmlichen Emotionsausdruck und -eindruck sind neben der Musterkennzeichnung auf merkmalsanalytische Beschreibungen angewiesen. Sie basieren meist auf Kommunikationsexperimenten mit speziell hergestellten Stimuli, z.B. Einzelwörtern, Sätzen, (seltener) Texten, die von professionellen Sprechern mit kontrolliert intendiertem Emotionsausdruck gesprochen werden (vgl. z.B. Zilliken 1991; Neuber 2002; Wendt 2007; Redecker 2008). Vergleichbare Untersuchungs-

designs wurden bereits von der frühen Ausdruckspsychologie entworfen, vgl. z.B. die oben angeführten physiognomischen Radioexperimente oder die Forschungen zur phonetischen Lautstilistik; methodisch verfeinert sind sie noch immer etabliert, neben der Sprechwissenschaft z.B. auch in Emotionspsychologie und Psycholinguistik (vgl. z.B. Tischer 1993; Scherer 1995 und 2003; Klasmeier 1999; Scherer/Banse/Wallbott 2001; Paeschke 2003). Der komplexe, multifaktoriell bestimmte Prozess des Emotionsausdrucks und seiner Verarbeitung wird auf möglichst wenige kontrollierbare Variablen reduziert. Ausgangspunkt hierfür ist die Alltagserfahrung, dass trotz der genannten Komplexität Kommunikationsbeteiligte ständig intuitiv Vorausagen über Wirkungen machen, Wirkungen und darauf bezogene Reaktionen im voraus abschätzen, sich also im Komplexgeschehen anhand einzelner Merkmale orientieren.

3.3.1 Emotionsforschung anhand von Einzelwörtern

Exemplarisch wird das „Gesprochene Wortkorpus für Untersuchungen zur auditiven Verarbeitung von Sprache und emotionaler Prosodie“ (WaSeP[®]) vorgestellt (Wendt 2007). Im Leibniz-Institut für Neurobiologie in Magdeburg wird dieses Korpus für Grundlagenforschungen zur Sprach- und Emotionsverarbeitung im Gehirn verwendet, z.B. mit Hilfe von Verhaltenstests und Methoden der nicht-invasiven Bildgebung (z.B. fMRI, EEG). Bei diesen Messmethoden muss auf den kommunikativen Kontext verzichtet werden, und die Probanden, sowohl Gesunde als auch Patienten mit neurologischen Störungen, sind in der Regel nicht im analytischen Hören trainiert.

Ausgehend von der These, dass der stimmlich-artikulatorische Ausdruck von Emotionen sehr robust ist und es „auch dem untrainierten Hörer möglich [ist], stimmliche Emotionsausdrücke ohne kommunikativen Kontext noch relativ sicher zu erkennen und zuzuordnen bzw. zu decodieren“ (Wendt 2007, S. 86), hat die Autorin ein Korpus aus zweisilbigen deutschen Substantiven und Pseudowörtern ohne kommunikativen Kontext erstellt. Diese Wörter wurden von je einer Schauspielerin und einem Schauspieler mit simuliertem Emotionsausdruck zu vorgegebenen Basisemotionen gesprochen (Freude, Trauer, Wut, Angst, Ekel, zum Kontrast: Sachlichkeit). Auf diese Weise sind Prototypen stimmlich-artikulatorischen Emotionsausdrucks entstanden. Zwar bilden reine Basisemotionen in der kommunikativen Realität eher die Ausnahme, für neurologische Untersuchungen ist diese begrenzte Vorgabe aber nach Wendt zugunsten der Überschaubarkeit für die Probanden und der Interpretierbarkeit der Ergebnisse hinnehmbar.

Das Wortkorpus ist mittels Hörerbefragungen (Identifizierung des intendierten Emotionsausdrucks), auditiven Expertenbeurteilungen und exemplarischen akustischen Beschreibungen (phonische Merkmale) validiert worden. Es liegen Konsensusbeurteilungen der stimmlich-artikulatorischen

Merkmale für die genannten Emotionsausdrücke vor, außerdem für ein Pseudowort in allen Emotionen auch akustische Messungen. In der Tendenz waren bestimmte akustische Merkmale für bestimmte emotionale Ausdrucksmuster zu finden, ebenso Übereinstimmungen zwischen akustischer und auditiver Beschreibung. Es liegt aber keine eindeutige Korrelation zwischen auditiven und akustischen Beurteilungen vor, die Ausdrucksmuster sind nicht ausschließlich durch Merkmalscluster charakterisiert. Es handelt sich um Gestalten, für deren Wahrnehmung weniger ihre einzelnen Komponenten als die Relationen zwischen ihnen ausschlaggebend sind. Nach akustischer Messung und auditiver Beurteilung kommen die Unterschiede zwischen den Sprechfassungen der einzelnen Basisemotionen nicht nur durch stimmliche Variation zustande, sondern beide Sprecher variieren systematisch auch zahlreiche andere sprecherische Merkmale (z.B. Lautheit, Sprechspannung, Artikulationspräzision, Lautdauer). Es handelt sich also um einen Komplex an Ausdrucksmerkmalen. Darüber hinaus weisen die Sprechfassungen trotz der kurzen Wort-Stimuli und der klaren Emotionsvorgaben individuelle Ausdrucksvarianten auf.

Ein Vorteil dieses standardisierten Korpus liegt neben der guten technischen Qualität der Audioaufnahmen in der Vergleichbarkeit der Sprechfassungen aufgrund der identischen Wortvorgaben. Solche strikt experimentell vorgehenden Untersuchungen können gerade durch die bewusste Abstraktion von der Komplexität der Kommunikationswirklichkeit Erkenntnisse über die (mögliche) Relevanz einzelner Variablen, über die Interdependenz von Merkmalsausprägung und Hörerbeurteilung und über die Entwicklung und Validierung analytischer Beschreibungskategorien liefern. Die Übertragbarkeit der Ergebnisse auf Bedingungen authentischer Kommunikation ist aber fraglich, denn hier kommt stimmlich-artikulatorischer Ausdruck reiner Basisemotionen ausgesprochen selten vor, typisch sind dagegen abgeschwächte Mischrealisationen in verschiedenen Schattierungen, die sich zudem abhängig vom Prozess der Kommunikation verändern. Im vorliegenden Wortkorpus sind dagegen die stimmlich-artikulatorischen Realisationen der vorgegebenen Emotionen aufgrund der sehr kurzen Wörter statisch. Dieses Material kann jedoch ebenso wie die parameterorientierten Beschreibungssysteme aus der klinischen Sprechstimmforschung für das Training im funktionell-analytischen Hören genutzt werden (ausführlich vgl. Bose/Schwarze/Wendt ersch. 2009).

3.3.2 Emotionsforschung anhand von quasi-authentischen Sprachäußerungen

Andere Studien verwenden als Stimuli umfangreicheres quasiauthentisches sprachliches Material und verbinden Analysen des Emotionsausdrucks mit experimentellen Untersuchungen zu seiner Wirkung. So haben Braun/

Heilmann (stellvertretend 2005) emotionale Ausdrucksmuster im synchronisierten Film untersucht. Anhand von Kurzszenen der Serie „Ally McBeal“ (amerikanische Originalfassung, deutsche und japanische Übersetzung) wurde die stimmliche Ausgestaltung einiger Basisemotionen (Angst, Trauer, Wut, Freude) phonetisch analysiert und ihre interkulturelle Erkennbarkeit und Akzeptanz (durch deutsche, amerikanische und japanische Hörer und Zuschauer) wurde anhand von Fragebögen mit vorgegebenen Emotionen geprüft. Für einige Basisemotionen (Trauer und kalte Wut) weisen die Autorinnen sowohl im Ausdruck als auch im Eindruck interkulturelle Asymmetrien nach (Braun/Heilmann 2005, S. 187).

Nebert (2007 und 2009) hat die sprachspezifische Verwendung der mittleren Sprechstimmlage für zwei Sprachen (russisch und deutsch) untersucht. Unter der mittleren Sprechstimmlage wird die wahrgenommene durchschnittliche Sprechtonhöhe verstanden, um die sich die Tonhöhenvariationen beim Sprechen bewegen. Sie wird u.a. durch emotionale Erregung beeinflusst, ist aber auch sprachspezifisch determiniert (vgl. Braun 1994 und 2001).

Deutsche Hörer bewerten Deutsch sprechende Russinnen oft als emotional, auch in sachlich intendierten Äußerungen, und geben als Grund deren hohe Sprechstimmlage an. Mittels akustischer Messungen weist Nebert (2007) jedoch nach, dass die mittlere Sprechstimmlage in beiden Sprachen sehr ähnlich ist – die russischen Sprecherinnen sprechen im Mittel lediglich einen Halbton höher. Dabei nutzen sie aber einen größeren Sprechtonumfang und die Tonhöhen sind ungleich verteilt: Im Vergleich zu den deutschen Sprecherinnen verwenden die russischen Sprecherinnen höhere Frequenzen häufiger. Sie sprechen bei gleicher Sprechstimmlage außerdem häufiger im Kopfregeister, wodurch ein hellerer Klang entsteht, und ihre Sprechmelodie steigt in akzentuierten Silben kurzzeitig an. Bei den deutschen Sprecherinnen ist dagegen eine Tendenz zum Brustregister festzustellen, verbunden mit eher dunklem Stimmklang. Deutsche Hörer interpretieren diese Stimmverwendungsmuster vermutlich deshalb als emotional, weil im Deutschen bereits geringe Abweichungen der Tonhöhe von einem Halbton an oberen Positionen der Sprechmelodie als emotionale Markierung gelten.

Nach Nebert (2009) übertragen die Sprecherinnen den Stimmverwendungsmodus der Muttersprache in die Fremdsprache: Russische Muttersprachlerinnen klagen oft über Stimmbeschwerden, weil sie die vermeintlich tiefere Lage des Deutschen auf unphysiologische Weise zu erreichen versuchen (Pressen, Drücken, Verlagern der Stimme). Im Gegensatz dazu versuchen deutsche Sprecherinnen, die vermeintlich höhere Lage des Russischen mit enger, gespannter Stimmgebung zu erreichen.

3.4 Perzeption von stimmlich-artikulatorischem Ausdruck

3.4.1 Sprechwirkungsforschung

Gegenstand der Sprechwirkungsforschung ist der beobachtbare bzw. messbare Effekt, den gesprochene Äußerungen bei Rezipienten auslösen (z.B. Krech/Richter/Stock/Suttner 1991; Hirschfeld/Neuber/Stock 2009). Untersucht werden vor allem Wirkungen stimmlich-artikulatorischer Ausdruckserscheinungen in ihrer Interdependenz mit sprachlichen und non-verbalen Mitteln. Wirkungen entstehen im Rezeptionsprozess; sie können zwar vom Sprecher intendiert sein, aber er hat keine Garantie für die Umsetzung seiner Intentionen. Denn Wirkungen hängen wesentlich davon ab, welche Motive ein Rezipient dem Sprecher unterstellt, wie er Inhalte und antizipierte Folgen der Äußerungen bewertet, welche Relevanz er der Situation zuschreibt (Hirschfeld/Neuber/Stock 2009, S. 773). Diese bewussten und unbewussten Bewertungen basieren auf subjektiven Antizipationen und Wahrnehmungsfiltren, die die Erwartungen an das Handeln des Partners und die Wahl der eigenen, auf den Partner zielenden Handlungen steuern (ebd., S. 779). Zugrunde gelegt wird ein prozessdynamisches Rezeptionsmodell der Integration von Kognition und Emotion, der Interdependenz sprechsprachlicher Komplexsignale mit nichtsprachlichen Umgebungsinformationen, gespeichertem deklarativem Wissen und Bewertungen der bereits vorliegenden Informationen (z.B. Herrmann 1995).

Sprechwirkungsforschungen zielen vor allem auf interindividuelle und vorhersagbare, systematische Gemeinsamkeiten im kommunikativen Effekt. Damit verbunden sind überwiegend experimentelle Versuchsanordnungen aus einer externen Perspektive (Variablenisolierung, Kommunikationsexperimente, Hörerbefragungen). Das hat zwar den Vorteil, dass ein Zugriff auf das komplexe Ursachen-Wirkungsgefüge relativ gut möglich ist, weil Auslösefaktoren und dadurch erzeugte Wirkungen kontrolliert werden können. Das hat jedoch auch den Nachteil, dass so kaum Erkenntnisse über die Funktionsweise der beobachteten Zusammenhänge im Kommunikationsprozess, also über die Interpretation der beobachteten Phänomene durch die Beteiligten im Kommunikationsprozess und ihre Relevanz für die gemeinsame Gesprächskonstitution zu gewinnen sind. Stellvertretend werden zwei Arbeiten knapp vorgestellt, die mit speziell hergestellten, systematisch variierten Sprachstimuli arbeiten, um den Einfluss bestimmter Variablen auf Rezipienten zu testen.

Neuber (2002) hat zwei Vorlesefassungen einer Erzählung – eindringlich versus monoton – phonetisch analysiert und bezüglich ihrer Wirkung (Fragebogen zur Bewertung von Text und Sprecher, Behaltenstest) an etwa 200 Hörern getestet. Er konnte nachweisen, dass stimmlich-artikulatorischer Ausdruck wesentlich für die Bildung globaler Textkohärenz verant-

wortlich ist: Die eindringliche Sprechfassung wurde nicht nur besser bewertet (als verständlicher, strukturierter, interessanter, emotionaler) als die monotone Fassung, sondern auch besser behalten (vor allem aufgrund von Melodie- und Tempogestaltung, Sprechgliederung und Pausen-Akzent-Struktur).

Redecker (2008) hat den Einfluss verschiedener stimmlich-artikulatorischer Parameter auf die Perception eines Werbevideos für Parfüm geprüft. Es wurden drei Varianten hergestellt: während Bilderfolge und Kurztext jeweils gleich blieben, variierte ein professioneller Sprecher die sprecherische Realisation (mit den Assoziationen eines ‚frischen und leichten‘, ‚athletischen und dynamischen‘, ‚würzigen und schweren‘ Duftes; vgl. Redecker 2008, S. 88 ff.). Die Sprechfassungen wurden auditiv und akustisch analysiert. Drei Probandengruppen (insgesamt etwa 150 Befragte) wurden mit jeweils einer Variante des Werbevideos konfrontiert und mit einem Fragebogen nach ihrem Eindruck befragt. Es zeigte sich, dass bereits geringe stimmlich-artikulatorische Veränderungen hochsignifikante Veränderungen in den Angaben zur Kaufabsicht und zur Bewertung von Werbespot, Duft und dargestellter Person zur Folge hatten.

3.4.2 Perception von pathologischem Fremdakzent

Im Folgenden wird eine Perzeptionsstudie vorgestellt, die in gewisser Weise an das literarische Beispiel von Edgar Allan Poe anknüpft, da auch hier Hörer durch ungewöhnlichen stimmlich-artikulatorischen Ausdruck irritiert sind und ihn fälschlich als fremdsprachlich interpretieren. Es geht um einen Fall von pathologischem Fremdakzent (Wendt/Bose/Sailer/Scheich/Ackermann 2007; Bose/Wendt ersch. 2010). Das sog. Fremdakzentsyndrom ist eine sehr seltene Sprach- und Sprechstörung, die z.B. nach einem Schlaganfall auftreten kann (vgl. stellvertretend Monrad-Krohn 1947; Blumstein/Kurowski 2006). Beobachtet werden immer verschiedene leichtere sprachliche und sprecherische Auffälligkeiten im Sprechen der Betroffenen, wodurch diese auf ihre Umgebung wie Sprecher einer Fremdsprache wirken.

Im vorliegenden Fall handelt es sich um eine 35-jährige Frau (AL) aus Sachsen-Anhalt, monolingual deutsch, mit Media-Infarkt links, hemiparetischen Störungen und motorischer Aphasie. Die Patientin berichtet, dass fremde Gesprächspartner sie z.B. beim Einkaufen als „russische Muttersprachlerin“ oder „Russlanddeutsche“ wahrnehmen würden. Sie selbst bemerke an sich diesen Fremdakzent nicht. Zusätzlich zu den vorhandenen körperlichen und kommunikativen Schwierigkeiten nach der Krankheit erlebe sie diese Erfahrung aber als Identitätsverlust:

abe:r wEnn ich mA! AUf dem mA:rkt bin und EIer KAufe:- |
 DANN:, (.) frAgen sie MICH- (.) ge!NAU!so. |
 WO:- (.) wo stAMmen sIE denn HE:R´ |
 [...]

und äh ich !mEr!ke das nicht dass ich so kOmisch SPREche; ||

Freies Sprechen und Vorlesen von AL weisen unspezifische Auffälligkeiten auf, sowohl in Lexik und Grammatik als auch in Artikulation und Prosodie. Die phonetischen Auffälligkeiten ergeben allerdings nicht das typische Gesamtbild eines russischen Akzents im Deutschen (vgl. hierzu u.a. Müller 2002; Stock/Velichkova 2002; Gorozhanina 2007), sondern erweisen sich als kaum systematisch. Sie liefern nur sehr eingeschränkte Hinweise darauf, warum fremde Kommunikationspartner die Patientin als russische Muttersprachlerin wahrnehmen. Es ist zu vermuten, dass sowohl segmentale als auch suprasegmentale Abweichungen dafür verantwortlich sind, insbesondere fehlende Reduktion unbetonter Endsilben, Sprechunflüssigkeit und atypische Sprechgliederung (vgl. ausführlich Wendt/Bose/Sailer/Scheich/Ackermann 2007 mit Klangbeispielen).

Für die Perzeptionsstudie wurden aus der Audioaufnahme eines Gesprächs zwischen der Patientin und einer Logopädin über Alltagsthemen sieben kurze Äußerungen von AL ausgewählt (Umfang 14 bis 37 Wörter). Diese Gesprächsausschnitte wurden bisher in Interviews 46 deutschsprachigen Nicht-Expert/inn/en unterschiedlicher regionaler Herkunft vorgespielt (vgl. Bose/Wendt ersch. 2010). Nach dem Hören jedes Ausschnitts wurden die Hörer/innen ohne inhaltliche Vorgabe um Assoziationen zum Gehörten gebeten. Hier einige Ergebnisse, zunächst zu den Angaben über die Sprecherin:

- Oft schon nach dem Hören des ersten oder zweiten Ausschnitts äußerten sich die Befragten zur Herkunft der Sprecherin. In 42 Interviews wurde die Sprecherin als „nichtdeutsch“ klassifiziert. Als Muttersprache wurde in 35 Interviews eine osteuropäische Sprache vermutet, in 25 Interviews speziell Russisch.
- In 35 Interviews wurde eine pathologische Ursache für die auffällige Sprechweise vermutet (Sprach-, Sprech-, Hörstörungen oder eine kognitive Störung), meist erst nach dem Hören der vierten oder fünften Äußerung.
- Die Befragten stellten Vermutungen an über die Stimmung der Sprecherin (z.B. *kräftlos/leidend/hoffnungslos*), über ihren Charakter und sozialen Status (z.B. *träge/langsam denkend/isoliert/bildungsfern*).

Viele Befragte machten darüber hinaus konkrete Angaben zu den gehörten Äußerungen:

- Die sprachliche Gestaltung wurde ebenso als auffällig beurteilt (z.B. *sehr einfacher Stil/unvollständige, abgebrochene Sätze/viele Wiederholungen*) wie der stimmlich-artikulatorische Ausdruck (z.B. *eigenartig/angestrengt/zögernd*).
- Beschrieben wurden aber auch zahlreiche konkrete stimmlich-artikulatorische Merkmale: z.B. Sprechtonhöhe/Melodie (*tief/monoton*), Stimmklang (*dumpf/dunkel/klagend*), Lautheit und Sprechspannung (*leise/monoton/lasch*), Sprechgeschwindigkeit (*schleppend/sehr langsam*), Rhythmus (*stockend/abgehackt*), Akzentuierung (*falsche Betonungen/gleichförmig*), Artikulation (*bemüht sehr deutlich/Laute und Silben entweder verschluckt oder langgezogen/nicht typisch deutsch*).

Angesichts dessen, dass hier kein authentischer Fremdakzent vorliegt und dass die phonetische Beschreibung kein einheitliches Bild ergibt, erstaunt die Schnelligkeit der Urteile ebenso wie die große Übereinstimmung in der Zuschreibung der fremdsprachlichen Herkunft. Als Anhaltspunkte für das Urteil „russische Sprecherin“ nannten die Befragten vor allem Abweichungen in der Artikulation (z.B. abweichende Lautbildungen, Dehnung äußerungsfinaler Silben), in der Prosodie (z.B. merkwürdiger Melodieverlauf, abweichende Äußerungsakzentuierung, fremder Sprechrhythmus), seltener sprachliche Auffälligkeiten (z.B. sehr kurze und simple Sätze).

Die Patientin bedient die erfahrungsbasierten Erwartungen fremder Kommunikationspartner an situationsangemessene Sprachverwendung und stimmlich-artikulatorischen Ausdruck nicht ausreichend. Ihre Abweichungen sind aber offensichtlich nicht groß genug, dass sie sofort als pathologisch diagnostiziert würden, deswegen verursachen sie den Eindruck von Fremdheit. Die sehr schnell entstehenden Eindrücke über die Sprechweise und die daraus resultierende Bewertung der Sprecherin haben die Wahrnehmung und Deutung der Informationen auf der verbalen Ebene gesteuert: Die meisten Hörer/innen haben die Sprecherin zunächst als „Ausländerin“ identifiziert und deshalb die wahrgenommenen sprachlichen Fehler und sehr einfachen sprachlichen Strukturen als Zeichen „guter Fremdsprachenbeherrschung“ interpretiert. Etwa die Hälfte der Befragten blieb bei der einmal getroffenen Diagnose, die anderen zeigten sich zunehmend irritiert, je mehr Ausschnitte sie hörten, und änderten ihre Diagnosen mehrfach. Offensichtlich war die Dauer der Hörwahrnehmung ausschlaggebend: Je mehr Äußerungen diese Hörer/innen analysieren konnten, umso unsicherer wurden sie speziell in der Einschätzung der Sprecherin als Ausländerin. Offensichtlich entdeckten sie nach und nach zu viele widersprüchliche verbale und phonische Hinweisreize, die nicht mit dem zunächst identifizierten Fremdsprachenakzent übereinstimmten, so dass sie die erste Interpretation durch eine andere ersetzten (z.B. Sprach-, Hör- oder Kognitionsstörung).

Die Befragung zeigt, dass Rezipienten, selbst wenn sie nur kurze Äußerungen einer unbekanntem Sprecherin gehört haben, nicht nur komplexe Zuschreibungen äußern (über die Herkunft, die Situation und die Stimmung der Sprecherin), sondern auch differenziert Auskunft zu einzelnen Merkmalen geben können. Sie nehmen dabei sowohl auf die sprachliche als auch auf die stimmlich-artikulatorische Ebene Bezug und nutzen beide für Diagnose und Identifikation von Sprecherzustand und -intention. D.h., Rezipienten bewerten Äußerungen und Sprecherin sowohl hinsichtlich einer individuellen Spezifik (Physiognomik), z.B. in der Zuschreibung als „Ausländerin“, als auch hinsichtlich einer situationsbezogenen Spezifik (Pathognomik), z.B. in der Zuschreibung von Einstellungen als „unmotiviert/unschlüssig“. Vermutlich handelt es sich um Mustererkennungsprozesse im Sinne der Gestaltpsychologie: Wenn Hörer auf ‚signifikante Hinweisreize‘ der verbalen oder phonischen Ebene stoßen, dann ergänzen sie die dazugehörige Gestalt, indem sie das aktuell Gehörte mit internen Mustern und situationsbezogenen Erfahrungen und Erwartungen vergleichen.

3.5 Stimmlich-artikulatorischer Ausdruck in der Interaktion

Im Folgenden werden Arbeiten vorgestellt, in denen mit einem gesprächsanalytischen Zugang stimmlich-artikulatorischer Ausdruck als kommunikative Ressource im Gespräch analysiert wird. Anhand von Aufzeichnungen authentischer Kommunikationsereignisse wird in gesprächsanalytischen Forschungen die Prozesshaftigkeit von Gesprächen berücksichtigt, Gesprächsordnung und wahrnehmbare systematische und meist routinisierte Gesprächspraktiken werden rekonstruiert (für die Methode vgl. z.B. Depermann 1999). Dazu gehören nicht nur sprachliche Verfahren – in gesprächsanalytischen Arbeiten wurde die Relevanz prosodischer wie auch die nonverbaler Mittel für die verbale Interaktion nachgewiesen (vgl. z.B. stellvertretend Selting 1995; Couper-Kuhlen/Selting 1996; Auer/Couper-Kuhlen/Müller 1999; Couper-Kuhlen/Reber in diesem Band; Schmidt/Depermann in diesem Band). Daran schließen sprechwissenschaftliche Studien zu Repertoires stimmlich-artikulatorischer Ausdrucksmuster in spielerischen und ernsthaften Interaktionsmodalitäten (Kallmeyer 1979) zwischen vertrauten Partnern an. Ihnen gemeinsam ist, dass der stimmlich-artikulatorische Ausdruck in engem Bezug zur verbalen Interaktionsebene in authentischen Gesprächen analysiert wird.

In der mündlichen Kommunikation ist grundsätzlich von polyrelationalen Form-Funktions-Beziehungen auszugehen, deswegen werden zu ihrer Beschreibung verschiedene Methoden miteinander verknüpft – sowohl naturwissenschaftlich messend-beurteilende Arbeitsformen aus Sprechwissenschaft und Phonetik als auch sozialwissenschaftlich rekonstruktiv-inter-

pretierende Arbeitsformen aus der Gesprächsforschung. Die Praktiken im Umgang mit stimmlich-artikulatorischem Ausdruck und ihre interaktive Relevanz für den Gesprächsverlauf werden rekonstruiert anhand wiederkehrender Phasen und Verfahren zur Gesprächssteuerung, anhand von Kommentaren über die Angemessenheit vs. Unangemessenheit der Gestaltung stimmlich-artikulatorischer Ausdrucksmuster und anhand von Interpretationsbedingungen und -regeln, denen die Beteiligten im Gespräch folgen. Höreindrücke über die verwendeten stimmlich-artikulatorischen Ausdrucksweisen werden mittels auditiver Beurteilungen durch Hörergruppen und exemplarischer instrumentalphonetischer Analysen intersubjektiviert. Stimmlich-artikulatorische Ausdrucksweisen werden sowohl holistisch interpretiert (Mustererkennung und -interpretation) als auch hinsichtlich ihrer Einzelmerkmale beschrieben.

Soziale Rollenspiele von Vorschulkindern

Aus einer empirischen Langzeitstudie zum stimmlich-artikulatorischen Ausdruck im sozialen Rollenspiel von Vorschulkindern (Bose 2003) geht hervor, dass Kinder stimmlich-artikulatorischen Ausdruck systematisch als kommunikative Ressource nutzen. Sie bauen in ihren Rollenspielen interaktiv Ausdrucksrepertoires auf, erwerben so miteinander stimmlich-artikulatorische Ausdruckskompetenz und differenzieren sie im Verlaufe ihrer Spielbiografie aus. Dabei orientieren sie sich an Vorbildern aus ihrer Kommunikationsumwelt, an frühen Eltern-Säuglings-Kommunikationen ebenso wie an Sprechrollenmustern und Praktiken vertrauter Erwachsener. Offensichtlich nehmen sie die jeweils individuellen stimmlich-artikulatorischen Ausführungen (z.B. von Bezugspersonen, aber auch in Institutionen und Medien) als konventionalisiert-musterhaft wahr, d.h. also als Praktiken der Performanz von Interaktionsrollen, Emotionen und Einstellungen, kategorisieren sie und wenden sie selbst spielerisch an. Die Untersuchung zeigt, dass Kinder im Vorschulalter bereits über die nötige Kompetenz verfügen, kulturell bedingte stimmlich-artikulatorische Ausdrucksmuster im Spiel so treffend zu gestalten, dass externe erwachsene Hörer sie wiedererkennen und sich an eigenes Kommunikationsverhalten erinnert fühlen. Dennoch gestalten die Kinder ihre spielerischen Kommunikationen nach eigenständigen Regeln bzw. Konventionen; es handelt sich um kindliche Typisierungen und Stilisierungen vorgefundener Muster in sehr expressiver Form. Insofern sind die gespielten Ausdrucksmuster eigene Schöpfungen der Kinder, was unter anderem in veränderten Situationsbezügen, in Hinzufügungen und Auslassungen von Ausdrucksformen deutlich wird.

Die beobachteten Kinderpaare gestalten mit musterhaften stimmlich-artikulatorischen Ausdrucksweisen Spielrollen und inszenieren fiktive Dialoge zwischen den Spielfiguren. Oft markieren sie allein durch eine pla-

kativ-kontrastierende, funktional eingesetzte stimmlich-artikulatorische Ausdrucksgestaltung, wer in welcher Haltung spricht, ohne das auch verbal zu erläutern (vgl. die Analysen in Bose 2001 und 2003 mit Klangbeispielen). Damit erweist sich die stimmlich-artikulatorische Ausdrucksform als unabdingbar für das Verständnis der entworfenen Spielsituation.

Mit Hilfe stimmlich-artikulatorischer Ausdrucksmuster stützen und strukturieren die Kinder zudem ihre Interaktion, indem sie den Spielrahmen aufrecht erhalten bzw. wieder herstellen, den Wechsel zwischen verschiedenen Spielebenen anzeigen oder bestimmte Aktionen als spielerisch bzw. ernsthaft markieren. Und sie gestalten ihre Beziehung zum Spielpartner, indem sie mittels Ausdruckskongruenz (Konsonanz, Iteration, Ergänzung) die besondere Nähe zum Partner und die Zustimmung zu seinen Spielhandlungen unterstreichen, inhaltliche Korrekturen zu den Partneräußerungen vorbringen, potenzielle Auseinandersetzungen verhindern oder abschwächen, aber auch Imageverletzungen deutlich machen.

Die Studie zeigt, dass kindliche Kommunikationskompetenz neben der Sprachhandlungskompetenz auch die Kompetenz im stimmlich-artikulatorischen Ausdruck einschließt. Hierunter werden überwiegend implizite Kenntnisse über und Fähigkeiten zur situations- und stimmungsabhängigen Variation von stimmlich-artikulatorischen Ausdrucksweisen verstanden, die als Regulationsbasis für kommunikatives Handeln dienen. Sie sind mit interiorisierten Erwartungsvorstellungen über die Adäquatheit konkreter Ausdrucksgestaltungen verbunden, haben also Bezug zur Kommunikations- und Kulturgemeinschaft, in der die Kinder aufwachsen. Die Analysen belegen die Verbindlichkeit, die Kinder bestimmten stimmlich-artikulatorischen Ausdrucksmustern für bestimmte Spielrollen zumessen. Aus der interaktiven Verwendung der untersuchten Ausdrucksmuster der Spielrollenidentität in entsprechenden fiktiven Situationskontexten ist auf eine implizite Musterkenntnis der beobachteten Vorschulkinder zu schließen. Diese zeigt sich unter anderem in praktizierten Formen (z.B. systematische Gestaltung stimmlich-artikulatorischen Ausdrucks in Abhängigkeit von realen sozialen Rollen und Situationen) und in gespielten Formen (z.B. Gestaltungen von Spielrollen oder Spielebenen in inszenierten fiktiven Spielrahmen). Aus Wiederholungen und Variationen verschiedener Ausdrucksgestaltungen durch den Spielpartner, oft begleitet von metakommunikativ kommentierendem Lachen, vor allem aber aus gelegentlichen metasprachlichen Benennungen und Bewertungen bestimmter Ausdrucksweisen wird deutlich, dass die Kinder auch ein explizites Musterwissen haben.

Freizeitinteraktionen vertrauter Jugendlicher

Auch vertraute Jugendliche nutzen in ihren Freizeitinteraktionen solche Repertoires ausgiebig und kreativ, um z.B. mittels Performance und Expres-

sivität Geselligkeit herzustellen (Schäfer 2003; vgl. die Zusammenfassung in Bose/Ehmer 2007). Vor allem geht es ihnen um Identifizierung und Gestaltung von sozialen und medialen Klischees oder Dialekten. Die Jugendlichen spielen sowohl mit dem stimmlich-artikulatorischen Ausdruck konkreter Personen als auch mit demjenigen sozialer Rollen und Muster und integrieren ihn ebenso wie die Kinder in neue, spielerische Kontexte. Sie greifen überdies auf ein Repertoire an kommunikativen Erfahrungen der Gruppe zurück (z.B. Handlungsstereotype) und erfinden stimmlich-artikulatorische Spiele (z.B. Beschimpfungsrituale). Interaktiv wird in mehreren Runden sowohl kooperativ als auch kompetitiv die beste sprecherische Gestaltung herausgearbeitet; die verschiedenen Gestaltungsversuche werden kommentiert, bewertet und korrigiert, besonderes gelungene Gestaltungen werden honoriert (z.B. durch Lachen). Der stimmlich-artikulatorische Ausdruck ist komplex mit dem Gesamtgeschehen verwoben und erfüllt vielfältige Funktionen (z.B. Themeninitiierung und -progression, Markierung von Kontext und Interaktionsmodalität, Ausdruck von Haltungen und Bewertungen).

Gesellige Tischgespräche vertrauter Erwachsener

Gesellige Tischgespräche vertrauter junger Erwachsener weisen spontane Inszenierungen bzw. gemeinsame Fiktionalisierungen (Kotthoff 1999) auf (Ehmer 2004; vgl. die Zusammenfassung in Bose/Ehmer 2007). Diese Gesprächspraktik ist mit den kindlichen Rollenspielen vergleichbar, denn die Sprecher übernehmen ebenfalls andere als ihre alltagsweltlichen Rollen und animieren fiktive Figuren. Damit geht ein Wechsel der Interaktionsmodalität in eine humorvolle, spielerische Modalität einher. Hergestellt wird diese Modalität sowohl durch Lachen als auch über einen Rahmenbruch, der durch die Etablierung von Innen- und Außenperspektive entsteht. Dabei überbieten sich die Gesprächsteilnehmer gegenseitig, indem sie immer neue Pointen setzen, die von den anderen durch Lachen honoriert werden. Die Beteiligten sprechen nicht nur über die Szene bzw. Spielwelt, sondern agieren in Spielidentitäten bzw. animieren Figuren. Wesentlich stärker als die spielenden Kinder nehmen die erwachsenen Sprecher eine ‚kritische Distanz‘ zu den animierten Figuren ein, was besonders in der Ausdifferenzierung von Innen- und Außenperspektive deutlich wird.

Der stimmlich-artikulatorische Ausdruck leistet einen wesentlichen Beitrag zur Strukturierung der konversationellen Praktik. So wird die Inszenierung mitunter über einen interaktiv hergestellten Rhythmus im Wechsel der Sprechbeiträge als gestalthafte Ganzheit vom umgebenden Gespräch abgesetzt. Ähnlich wie in den kindlichen Rollenspielen trägt der stimmlich-artikulatorische Ausdruck auch dazu bei, zwischen dem Sprechen in der alltagsweltlichen Identität und dem Sprechen als Figur zu unterscheiden.

Wie schon die untersuchten Kinder und Jugendlichen übernehmen die erwachsenen Sprecher Ausdrucksgestaltungen ihrer Interaktionspartner, variieren und pointieren sie in spielerischem Wettstreit.

Fernsehkommunikation in Familien

Ähnliche Phänomene zeigt Baldauf (1998 und 2002) für fernsehbegleitende Kommunikation in Familien auf. Es handelt sich um offene Sprechsituationen (Goffman 1981), in denen die Zuschauer sowohl auf das im Fernsehen Gezeigte reagieren als auch untereinander darüber kommunizieren können. Auch hier stellen die Beteiligten oft eine spielerische, unernste Interaktionsmodalität her, überwiegend mit sprachlich knappen, pointierten Äußerungen als Kommentar zur laufenden Sendung. Gerade die verbale Kürze und Knappheit provoziert nach Baldauf (2002) ‚phonische Ausführlichkeit‘, üppig ausgearbeitete spielerische stimmlich-artikulatorische Ausdrucksformen. Eine große Rolle spielen sog. ‚minimale körpernahe Äußerungsformen‘ wie z.B. Lachen, Stöhnen, Seufzen oder Aufschreien, die Baldauf (1998, S. 40) als ‚lautlich, aber nicht notwendig sprachlich‘ definiert. In ihnen wird die Körperlichkeit des Sprechens offenbar und die Beteiligten stellen Informalität und Vertrautheit der Kommunikation heraus. Die Autorin weist nach, dass es sich bei diesen körpernahen Äußerungsformen nicht ausschließlich um individuelle, unwillkürliche Affektausdrücke handelt, sondern dass sie gleichzeitig auch eine kommunikative Funktion haben: Sie werden in den sprachlichen Kontext sequenziell eingebunden, richten sich auch an die anderen Familienmitglieder – und sie werden von diesen aufgegriffen und gemeinsam verarbeitet.

Konfliktäre Gespräche vertrauter Partnerinnen

Auch in ernsthaften Interaktionsmodalitäten werden komplexe Handlungen mittels stimmlich-artikulatorischer Ausdrucksgestaltungen lokal charakterisiert. Schwarze (2009) zeigt das in einer Studie zu Formen und Funktionen der Topoiverwendung in konfliktären Gesprächen zwischen Müttern und ihren jugendlichen Töchtern. Hier trägt der stimmlich-artikulatorische Ausdruck nicht vordergründig zur Charakterisierung von performativen Praktiken oder Interaktionsrollen bei, sondern dient zur Kennzeichnung bestimmter Toposverwendungsweisen. Diese Verwendungsweisen unterscheiden sich im Hinblick auf die Interaktivität bei der Herstellung des Topos, also ob ein Topos von einer Partnerin allein oder von beiden Partnerinnen gemeinsam hergestellt wird. Topoi treten nach Schwarze (ebd.) vorzugsweise in der einen oder anderen Herstellungsweise auf. Weitere Unterscheidungskriterien liefern die Ebenen der Gesprächsorganisation, der toposbezogenen interaktiven Konsequenzen sowie der Äußerungsge-

staltung und Formulierungsdynamik (zu den Kategorien dieser Analyseebenen vgl. Deppermann 1999, S. 56 ff.).

Die Verschränkung von stimmlich-artikulatorischer Ebene, Äußerungsgestaltung und Formulierungsdynamik wird in den Analysen beispielsweise an prosodischen Parametern wie Sprechgeschwindigkeit, Sprechgliederung, Akzentposition und -häufigkeit deutlich, die Aufschluss über die Sprechplanung und über die Formelhaftigkeit bzw. Vorgeformtheit der Äußerungen geben (vgl. u.a. Bose 1994 und 2003). Insbesondere Akzentuierung und sprecherische Gliederung ermöglichen Inferenzen und können in der Argumentationsanalyse für die Rekonstruktion von Prämissen und Argumentstrukturen genutzt werden (vgl. Bose/Gutenberg 2003; Schwarze 2009). Es handelt sich bei den prosodischen Realisationen also nicht um Oberflächenerscheinungen, sondern sie sind Ergebnis eines Entscheidungsprozesses darüber, was nach Annahme der Sprecherin für ihre Partnerin in der konkreten Gesprächssituation besonders wichtig bzw. unwichtig ist.

In den auditiven Analysen der Konfliktgespräche fielen stimmlich-artikulatorische Gestalten auf, die auf der Handlungsebene mit einem individuell hergestellten Topos kookkurieren, außerdem mit flüssiger, ungestörter Formulierungsdynamik und mit Formelhaftigkeit auf der lexiko-semantischen Ebene. Diese Formelhaftigkeit (Phraseologismen, Stereotype usw.) weist auf allgemeine Topoi bzw. Gemeinplätze (*loci communes*) hin; nach Schwarze Anzeichen für eine ‚Orientierung am Modell‘ (Dausendschön-Gay/Gülich/Krafft 2007, S. 184), also für „eine diskursive Routine, bei der ein in spezieller, teils konventionalisierter sprachlicher Form etabliertes Modell in der laufenden Interaktion Bestand hat und von den Interagierenden aufgegriffen wird.“ (Schwarze 2009, S. 130). Die damit einhergehende stimmlich-artikulatorische Gestalthaftigkeit ist allerdings keine obligatorische, sondern eine fakultative Eigenschaft der Toposverwendung. Schwarze wertet sie als Bestandteil rhetorischer Verfahren, als phonische Kennzeichnung besonderer elocutioneller Gestaltung im Sinne einer *Amplificatio*.

Diese enge Verwobenheit der rhetorischen, argumentativ-topischen und ausdrucksbezogenen Gesprächsarbeit in Konfliktgesprächen hat, anders als in der Modalität der spielerischen Geselligkeit, keine unterhaltende und illustrierende Funktion. Sie dient vielmehr der Herstellung von Kohärenz und somit der Erkennbarkeit der Argumentationssequenz, damit spezifisch argumentationsbezogenen individuellen Zielen und Zwecken der Gesprächsteilnehmerinnen im Gespräch.

Die vorgestellten gesprächsanalytischen Studien zeigen, dass vertraute Kommunikationspartner den stimmlich-artikulatorischen Ausdruck als kommunikative Ressource nutzen, um verschiedene Gesprächspraktiken (kindliches Rollenspiel, Beschimpfungsrituale, spontane Inszenierungen, konversationelle Karikaturen, aber auch konfliktäre Argumentationen) zu

organisieren. Zum einen gestalten Sprecher ihre Äußerungen entsprechend konventioneller und damit wiedererkennbarer Muster, wodurch relevantes Wissen auf verschiedenen Ebenen verfügbar gemacht werden kann (soziale Konnotationen, Sprechrollen, Gestalthaftigkeit, Toposverwendung usw.). Zum anderen wird der stimmlich-artikulatorische Ausdruck von den Sprechern variiert. Die Variabilität bietet den Gesprächsteilnehmern die Möglichkeit, sowohl Gemeinsamkeit zu schaffen als auch mit den anderen zu kompetitieren.

4. Zusammenfassung

Die in neueren geisteswissenschaftlichen Arbeiten diskutierte Vielschichtigkeit bzw. Komplexität des Phänomens Stimme spiegelt sich in den hier dargestellten theoretischen Modellvorstellungen und Methoden zur empirischen Beschreibung von stimmlich-artikulatorischem Ausdruck wider. In ihnen werden unterschiedliche Aspekte der Sprechstimme fokussiert – ihre Physiologie (Stimme als Organ) und ihre kommunikative Funktion und soziale Determiniertheit (Stimme als Bestandteil des stimmlich-artikulatorischen Ausdrucks, beschreibbar in engem Bezug zur Sprache als individuelle Ausdrucksweise oder als konventionalisiertes Muster). Entsprechend wurden in den ausgewählten empirischen Studien sowohl stimmlich-artikulatorische Merkmale beschrieben und klassifiziert (Stimmphysiologie-, Emotionsforschung) als auch stimmlich-artikulatorischer Ausdruck in seiner Wirkung auf Hörer (Sprechwirkungs-, Perzeptionsforschung) und in seiner interaktiven Ausarbeitung (Gesprächsforschung) betrachtet.

Stimmlich-artikulatorischer Ausdruck dient zur Charakterisierung von Sprecher und Situation, von Zugehörigkeit zu bestimmten sozialen oder regionalen Gruppen u.a.m.; er verweist nicht nur auf das Individuum, das seinen Empfindungen willkürlich oder unwillkürlich Ausdruck gibt, sondern auch auf den Adressaten, bei dem Empfindungen und Assoziationen angeregt werden, und damit auf den Interaktionsprozess. Rezipienten deuten einen konkreten stimmlich-artikulatorischen Ausdruck vor dem Hintergrund ihrer Erfahrungen bzw. Erwartungen über angemessenes Ausdrucksverhalten als Ausdruck von Persönlichkeitseigenschaften, von Emotionen und Einstellungen, von Interaktionsrollen. Ein Problem der Beschreibung des stimmlich-artikulatorischen Ausdrucks besteht darin, dass nicht alle Merkmale für eine konkrete Situation relevant sein müssen. Kommunikationsbeteiligte ziehen ihre Schlüsse darüber, wie etwas, was der andere sagt bzw. meint, zu verstehen und zu interpretieren ist, aufgrund von Hinweisreizen, Kontexthilfen und Situationsindizien, zu denen auch der stimmlich-artikulatorische Ausdruck gehört. Allerdings sind diese Hinweisreize nicht zweifelsfrei zu identifizieren oder vorherzusagen.

Aus dem Überblick über empirische Arbeiten zum stimmlich-artikulatorischen Ausdruck geht hervor, dass validierte parameterbasierte Beschreibungssysteme zur Verfügung stehen, die auf physiologischen und phonetischen Kriterien basieren (vgl. 2.3 und 3.2). Gemeinsam ist ihnen die Orientierung an klar definierten, transparenten Bewertungsmaßstäben und die Konzentration auf wenige wesentliche Klangphänomene der Stimme, die meist skaliert beschrieben werden. Dennoch kommen Beschreibung und Interpretation der einzelnen Parameter oft aufgrund komplexer, nicht immer genau aufzuschlüsselnder Höreindrücke zustande, und sie sind kulturabhängig (3.2.3). Der Zusammenhang zwischen akustischen Grundlagen und Hörempfindungen ist nicht linear, und es ist nicht klar zu trennen zwischen ausschließlich stimmlichen, prosodischen oder artikulatorischen Gestaltungsweisen. Vor allem aber ist der Komplex des stimmlich-artikulatorischen Ausdrucks anhand seiner physiologisch bestimmbaren Merkmale allein nicht adäquat zu erfassen.

In vielen empirischen Arbeiten wird der multifaktoriell bestimmte Prozess des stimmlich-artikulatorischen Ausdrucks und seiner Verarbeitung auf möglichst wenige kontrollierbare Variablen reduziert, indem experimentell erstelltes Sprachmaterial in kontrollierten Situationen analysiert und/oder in seiner Hörerwirkung getestet wird (vgl. 3.3 und 3.4.1). Dahinter steht die These, dass die untersuchten Merkmale in der Bündelung des Ausdrucksmuster zwar vielleicht nicht vollkommen erfassen, aber doch hinreichend beschreiben. Allerdings bestehen komplizierte Relationen zwischen Merkmalen und Mustern, die auf anthropologisch-evolutionärer Grundlage stark kulturell und sprachlich überformt und kontextuell abgestimmt sind. Strikt experimentell vorgehende Untersuchungen können deswegen zwar nachweisen, dass stimmlich-artikulatorische Ausdrucksformen als Bedeutungsträger fungieren und zur Vereindeutigung der Verständigung beitragen können; sie erbringen Erkenntnisse über die (mögliche) Relevanz einzelner Variablen, über die Interdependenz von Merkmalsausprägung und Hörerbeurteilung und tragen zur Entwicklung und Validierung analytischer Beschreibungskategorien bei. Die Übertragbarkeit der Ergebnisse auf die interaktive Relevanz des stimmlich-artikulatorischen Ausdrucks im Kommunikationsprozess ist aber nur sehr eingeschränkt möglich.

In der mündlichen Kommunikation ist nicht mit konstanten Merkmalskombinationen für bestimmte Funktionen zu rechnen: Kommunizierende können bestimmte Funktionen mit vielen verschiedenen sprecherischen Mitteln gestalten, die sie verschieden miteinander kombinieren. Dadurch ergeben sich eine hohe Varianz und eine prinzipielle Mehrdimensionalität der Ausdrucksformen in wechselnder Hierarchie und Kombination und eine Multifunktionalität von Mustern und Merkmalen. Diesen komplexen

Zusammenhang versuchen Perzeptionsstudien anhand von Interviews über Assoziationen zu authentischen Äußerungen zu erfassen (vgl. 3.4.2). Daraus geht hervor, dass der stimmlich-artikulatorische Ausdruck das Verständnis des Gesagten entscheidend beeinflusst, weil Rezipienten ihm außerordentlich schnell Indizien für die Diagnose und Identifikation von Sprecherpersönlichkeit und -intention entnehmen. Außerdem zeigt sich, dass Hörer nicht nur einen Gesamteindruck haben, sondern das Gehörte auch gezielt analysieren können, wobei sie gleichermaßen sowohl auf die sprachliche als auch auf die stimmlich-artikulatorische Ebene Bezug nehmen.

In gesprächsanalytischen Arbeiten werden die kommunikativen Praktiken im Umgang mit stimmlich-artikulatorischem Ausdruck und seine interaktive Relevanz in engem Bezug zur verbalen Interaktionsebene in authentischen Gesprächen untersucht (vgl. 3.5). Es zeigt sich, dass Gesprächsteilnehmer die Relevanz von stimmlich-artikulatorischen Ausdrucksformen interaktiv aushandeln und dass der stimmlich-artikulatorische Ausdruck komplex mit dem Gesamtgeschehen verwoben ist und vielfältige Funktionen erfüllt (z.B. Themeninitiierung und -progression, Markierung von Kontext und Interaktionsmodalität, Ausdruck von Haltungen und Bewertungen, Inszenierung von Sprechrollenmustern, Wiedergabe fremder stimmlich-artikulatorischer Ausdrucksmuster).

Hier bietet sich eine Verknüpfung der Arbeiten zum stimmlich-artikulatorischen Ausdruck mit den Konzepten der Kontextualisierung und der Performanz an. Methodisch arbeiten die vorgestellten Studien mit einer Kombination erstens aus rekonstruktiven Analysen von Gesprächsverlauf und -steuerung, zweitens aus externen (kulturbezogenen) Musterkennzeichnungen und drittens aus phonetischen Merkmalsbeschreibungen. Durch diese systematischen Methoden- und Perspektivenwechsel können die jeweiligen Interpretationen wechselseitig validiert werden.

Um den Komplex von stimmlich-artikulatorischen Ausdrucksformen in Interdependenz zu sprachlichen und körperlichen Ausdrucksformen umfassend theoretisch und empirisch zu erforschen, sind weitere, vor allem interdisziplinäre Forschungen nötig. Insbesondere sind die zahlreichen geisteswissenschaftlichen Perspektiven, aus denen Stimme und stimmlich-artikulatorischer Ausdruck aktuell thematisiert wird, zu berücksichtigen und miteinander in Beziehung zu bringen – durchaus im Sinne einer Forderung von Bühler (1933 [1982], S. 111): die Vertreter der beteiligten Disziplinen mögen „den Mut aufbringen, jeder dem anderen ins Konzept zu sprechen“.

Literatur

- Aderhold, Egon (1993): Sprecherziehung des Schauspielers. Grundlagen und Methoden. 4. Aufl. Berlin.
- The American Speech-Language-Hearing Association (ASHA) (2002): Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). Internet: www.asha.org/NR/rdonlyres/3FA67246-279B-4DA2-84D8-BEFC5D99345/0/22559_1.pdf (Stand: Juli 2009).
- Anders, Lutz Christian (1997): Spektrale Analysen gestörter Stimmen. Habilitationsschrift Universität Jena. Unveröff. Typoskript.
- Auer, Peter/Couper-Kuhlen, Elizabeth/Müller, Frank (1999): Language in time: The rhythm and tempo of spoken interaction. New York/Oxford.
- Bachtin, Michail (1972 [1981]): Slovo v romane. In: *Voprosy Literatury* 6/1972, S. 54–86. [Engl. Übers.: Discourse in the novel. In: Holquist, Michael (Hg.): *The dialogic imagination. Four essays*. Austin 1981, S. 259–422].
- Baldauf, Heike (1998): Aufschreien und Stöhnen – Äußerungsformen emotionaler Beteiligung beim Fernsehen. In: Brock, Alexander/Hartung, Martin (Hg.): *Neuere Entwicklungen in der Gesprächsforschung: Vorträge der 3. Arbeitstagung des Pragmatischen Kolloquiums Freiburg*. (= ScriptOralia 108). Tübingen, S. 37–54.
- Baldauf, Heike (2002): *Knappes Sprechen*. (= Reihe germanistische Linguistik 227). Tübingen.
- Blumstein, Sheila/Kurowski, Kathleen (2006): The FAS: a perspective. In: *Neurolinguistics* 19, S. 346–355.
- Bogner, Andrea (2000): Stimmen hören. Das Phänomen der Stimme in der interkulturellen Kommunikation. In: Wierlacher, Alois (Hg.): *Kulturthema Kommunikation. Konzepte – Inhalte – Funktionen*. Möhnesee, S. 209–217.
- Bogner, Andrea (2005): Grenzgänge. Zur konstitutiven Fremdheit der Stimme. In: *Jahrbuch Deutsch als Fremdsprache* 31, S. 108–130.
- Bose, Ines (1994): Zur temporalen Struktur frei gesprochenen Textes. (= *Forum Phonetikum* 5). Frankfurt a.M.
- Bose, Ines (2001): Methoden der Sprechausdrucksbeschreibung am Beispiel kindlicher Spielkommunikation. In: *Gesprächsforschung – Onlinezeitschrift zur verbalen Interaktion* 2, S. 262–303. Internet: www.gespraechsforschung-ozs.de (Stand: Juli 2009).
- Bose, Ines (2003): *dóch da sín ja ' nur müster // – Kindlicher Sprechdruck im sozialen Rollenspiel*. (= *Hallesche Schriften zur Sprechwissenschaft und Phonetik* 9). Frankfurt a.M. u.a.
- Bose, Ines/Ehmer, Oliver (2007): Formen und Funktionen des Sprechdrucks in geselliger Kommunikation. In: Velichkova, Ludmila (Hg.): *Klangsprache im Fremdsprachenunterricht (IV) – Forschung und Praxis*. Woronesh, S. 33–49.
- Bose, Ines/Gutenberg, Norbert (2003): Enthymeme and prosody. A contribution to empirical research in the analysis of intonation as well as argumentation. In: van Eemeren, Frans H./Blair, J. Anthony/Willard, Charles Arthur (Hg.): *Proceedings of the 5th International Conference on Argumentation (ISSA)*, Amsterdam June 25–28, 2002. Amsterdam, S. 139–140.

- Bose, Ines/Schwarze, Cordula/Wendt, Beate (ersch. 2009): Transkription von Sprechausdruck. In: *Gesprächsforschung – Onlinezeitschrift zur verbalen Interaktion*. Internet: www.gespraechsforschung-ozs.de (Stand: August 2009).
- Bose, Ines/Wendt, Beate (ersch. 2010): Sprechwissenschaftliche Studien zu einem Fall von Fremdakzentsyndrom (russischer Fremdakzent im Deutschen). In: Velichkova, Ludmila (Hg.): *Klangsprache im Fremdsprachenunterricht (VI) – Forschung und Praxis*. Woronesh.
- Braun, Angelika (1994): Sprechstimmlage und Muttersprache. In: *Zeitschrift für Dialektologie und Linguistik* 1/61, S. 170–178.
- Braun, Angelika (2001): Is voice quality language-dependent? In: Braun, Angelika/Masthoff, Herbert R. (Hg.): *Phonetics and its applications*. Stuttgart, S. 453–463.
- Braun, Angelika/Heilmann, Christa (2005): Fremde Stimmen und fremde Körper. Zum Ausdruck von Emotionen im synchronisierten Film. In: *Jahrbuch Deutsch als Fremdsprache* 31, S. 164–189.
- Bühler, Karl (1927 [1978]): *Die Krise der Psychologie*. Jena. [Ungek. Ausg. 1978. Stuttgart/Frankfurt a.M./Berlin/Wien].
- Bühler, Karl (1933 [1968]): *Ausdruckstheorie. Das System an der Geschichte aufgezeigt*. Jena. [2., unveränd. Aufl. 1968. Stuttgart/New York].
- Bühler, Karl (1934 [1982]): *Sprachtheorie: die Darstellungsfunktion der Sprache*. Jena. [Ungek. Neudr. d. Ausg. Jena 1934. Mit einem Geleitwort von Friedrich Kainz. 1982. Stuttgart/New York].
- Chion, Michel (2003): Mabuse – Magie und Kräfte des „Acousmètre“. Auszüge aus „Die Stimme im Kino“. In: Epping-Jäger, Cornelia/Linz, Erika (Hg.): *Medien/Stimmen*. Köln, S. 124–159.
- Connor, Steven (2004): „The Strains of the Voice“. In: Felderer (Hg.), S. 158–172.
- Couper-Kuhlen, Elizabeth/Selting, Margret (Hg.) (1996): *Prosody in conversation: interactional studies*. Cambridge.
- Couper-Kuhlen, Elizabeth/Reber, Elisabeth (i.d.Bd.): *Interjektionen zwischen Lexikon und Vokalität: Lexem oder Lautobjekt?*
- Dausendschön-Gay, Ulrich/Gülich, Elisabeth/Krafft, Ulrich (2007): Vorgeformtheit als Ressource im konversationellen Formulierungs- und Verständigungsprozess. In: Hausendorf, Heiko (Hg.): *Gespräch als Prozess. Linguistische Aspekte der Zeitlichkeit verbaler Interaktion. (= Studien zur deutschen Sprache 37)*. Tübingen, S. 181–219.
- Deppermann, Arnulf (1999): *Gespräche analysieren. Eine Einführung in konversationsanalytische Methoden. (= Qualitative Sozialforschung 3)*. Opladen.
- Dolar, Mladen (2004): Sechs Lektionen über Stimme und Bedeutung. In: Felderer (Hg.), S. 199–222.
- Drumbl, Johann (2005): Sprechen in vielen Stimmen als Aufgabe des Sprachunterrichts. In: *Jahrbuch Deutsch als Fremdsprache* 31, S. 92–107.
- Ehlich, Konrad (2004): Karl Bühler – zwischen Zeichen und Handlung oder: von den Mühen des Entdeckens und seinen Folgen. In: Ehlich/Meng (Hg.), S. 273–292.

- Ehlich, Konrad/Meng, Katharina (Hg.) (2004): Die Aktualität des Verdrängten. Studien zur Geschichte der Sprachwissenschaft im 20. Jahrhundert. (= Studien zur Wissenschafts- und Universitätsgeschichte 4). Heidelberg.
- Ehmer, Oliver (2004): Spontane Inszenierungen in Alltagsgesprächen. Unveröff. Dipl. arbeit. Univ. Halle a.d.S.
- Esling, John H. (2002): Laryngoscopic analysis of tibetan chanting modes and their relationship to register in sino-tibetan. In: Proceedings of the 7th International Conference on Spoken Language Processing (ICSLP), Bd. 2. Denver, S. 1081–1084.
- Evans, Ruth/Nawka, Tadeus (2007): Skalen für die auditive Stimmbewertung. In: Bose, Ines (Hg.): Sprechwissenschaft. 100 Jahre Fachgeschichte an der Universität Halle. (= Hallesche Schriften zur Sprechwissenschaft und Phonetik 22). Frankfurt a.M. u.a., S. 313–318.
- Fährmann, Rudolf (1960): Die Deutung des Sprechausdrucks. Studien zur Einführung in die Praxis der charakterologischen Stimm- und Sprechanalyse. Bonn.
- Felderer, Brigitte (Hg.) (2004): Phonorama. Eine Kulturgeschichte der Stimme als Medium. Berlin.
- Fitch, Tecumseh (2004): Die Stimme – aus biologischer Sicht. In: Felderer (Hg.), S. 86–102.
- Gauß, Eva Maria (2008): Sprechbildung für Schauspieler/innen. Das Praxis-konzept von Tara McAllister-Viel im Kontext der Diskurse zu kulturübergreifendem Theater und zu Stimmlichkeit. Unveröff. Dipl.-Arb. Univ. Halle-Wittenberg.
- Gauß, Eva Maria (2009): Welche Fragen wirft eine interkulturelle Mikroanalyse von Sprechbildungskonzepten für die Sprechwissenschaft auf? In: Hirschfeld, Ursula/Neuber, Baldur (Hg.): Aktuelle Forschungsthemen der Sprechwissenschaft: Phonetik, Rhetorik und Sprechkunst. (= Hallesche Schriften zur Sprechwissenschaft und Phonetik). Frankfurt a.M. u.a., S. 31–50.
- Geißner, Hellmut (1981): Sprechwissenschaft. Theorie der mündlichen Kommunikation. Königstein.
- Geißner, Hellmut (1984): Über Hörmuster. Gerold Ungeheuer zum Gedenken. In: Gutenberg, Norbert (Hg.): Hören und Beurteilen: Gegenstand und Methode in Sprechwissenschaft, Sprecherziehung, Phonetik, Linguistik und Literaturwissenschaft. (= Sprache und Sprechen 2). Frankfurt a.M., S. 13–56.
- Görlitz, Dietmar (1972): Ergebnis und Probleme der ausdruckspsychologischen Sprechstimmforschung. Meisenheim.
- Goffman, Erving (1981): Forms of talk. Philadelphia.
- Gorozhanina, Natalja (2007): Deutsch-russische rhythmische Interferenzen. In: Zeitschrift für Interkulturellen Fremdsprachenunterricht 12, 2. Internet: <http://zif.spz.tu-darmstadt.de/jg-12-2/navigation/startbei.htm> (Stand: August 2009).
- Grawunder, Sven (2009): On the physiology of voice production in South-Siberian throat singing – analysis of acoustic and electrophysiological evidence. Berlin.
- Gutenberg, Norbert (1998): Einzelstudien zu Sprechwissenschaft und Sprecherziehung. Arbeiten in Teilfeldern. Göppingen.

- Gutenberg, Norbert (2001): Einführung in Sprechwissenschaft und Sprecherziehung. Frankfurt a.M. u.a.
- Hartel, Gaby/Kaspar, Frank (2004): Die Welt und das geschlossene Kästchen: Stimmen aus dem Radio – und über das Radio. In: Felderer (Hg.), S. 132–144.
- Heilmann, Christa (2001): Neue Ansätze sprechwissenschaftlicher Gesprächsforschung. Habilitationsschrift Universität Jena.
- Herrmann, Theo (1995): Allgemeine Sprachpsychologie: Grundlagen und Probleme. 2. Aufl. Weinheim.
- Herzog, Herta (1933): Stimme und Persönlichkeit. In: Zeitschrift für Psychologie 130, S. 300–369.
- Hirano, Minoru (1981): Clinical examination of voice. (= Disorders of human communication 5). Wien/New York.
- Hirschfeld, Ursula/Neuber, Baldur/Stock, Eberhard (2009): Sprach- und Sprechwirkungsforschung. In: Fix, Ulla/Gardt, Andres/Knape, Joachim (Hg.): Rhetorik und Stilistik. Ein internationales Handbuch historischer und systematischer Forschung. 1. Halbbd. Berlin/New York, S. 772–786.
- Hirschfeld, Ursula/Stock, Eberhard (ersch. 2009): Gedanken Karl Bühlers zur Phonetik und Phonologie. In: Bachmann, Armin R./El Mogharbel, Christliebe/Himstedt, Katja (Hg.): Form und Struktur in der Sprache. Festschrift für Elmar Ternes. (= Tübinger Beiträge zur Linguistik 499). Tübingen.
- Höffe, Wilhelm (1965): Gesprochene Sprache. Ratingen.
- Kallmeyer, Werner (1979): (Expressif) eh ben dis donc, hein ‚pas bien‘ – Zur Beschreibung von Exaltation als Interaktionsmodalität. In: Klopfer, Rolf/Rothe, Arnold/Krauß, Henning/Kotschi, Thomas (Hg.): Bildung und Ausbildung in der Romania. München, S. 549–568.
- Kent, Raymond D./Ball, Martin J. (Hg.) (2000): Voice quality measurement. San Diego.
- Ketzmerick, Bettina (2007): Zur auditiven und apparativen Charakterisierung von Stimmen. (= Studententexte zur Sprachkommunikation 47). Dresden.
- Kirchhoff, Robert (1965): Grundfragen der Ausdruckspsychologie. In: Kirchhoff, Robert (Hg.): Handbuch der Psychologie, Bd. 5: Ausdruckspsychologie. Göttingen, S. 117–219.
- Kirchhoff, Robert (1991): Ausdruck. In: Arnold, Wilhelm/Eysenck, Hans Jürgen/Meili, Richard (Hg.): Lexikon der Psychologie. 3. Aufl. Freiburg/Basel/Wien, S. 185–194.
- Klasmeyer, Gudrun (1999): Akustische Korrelate des stimmlich emotionalen Ausdrucks in der Lautsprache. (= Forum Phonetikum 67). Berlin/Offenbach.
- Klawitter, Klaus/Köhler, Kerstin (2004): Gestisches Sprechen. In: Köhler, Kerstin/Skorupinski, Cäcilie (Hg.): Wissenschaft macht Schule. Sprechwissenschaft im Spiegel von 10 Jahren Sommerschule der DGSS. (= Sprechen & Verstehen 21). St. Ingbert, S. 151–162.
- Klawitter, Klaus/Minnich, Herbert/Honigmann, Ingeborg/Mont'agh, Imre (1981): Zur Sprecherziehung für Schauspieler. Berlin.

- König, Ekkehard/Brandt, Johannes G. (2006): Die Stimme – Charakterisierung aus linguistischer Perspektive. In: Kolesch/Krämer (Hg.), S. 111–129.
- Kolesch, Doris (2005): Wilde Laute und mediale Resonanzen. Fremde Stimmen im Theater der Gegenwart. In: Jahrbuch Deutsch als Fremdsprache 31, S. 208–215.
- Kolesch, Doris (2006a): Zehn Thesen über die Stimme. In: Bischoff, Doerte/Wagner-Engelhaaf, Martina (Hg.): Mitsprache, Rederecht, Stimmgewalt. Genderkritische Strategien und Transformationen der Rhetorik. Heidelberg, S. 341–356.
- Kolesch, Doris (2006b): Wer sehen will, muss hören. Stimmlichkeit und Visualität in der Gegenwartskunst. In: Kolesch/Krämer (Hg.), S. 40–64.
- Kolesch, Doris/Krämer, Sibylle (2006): Stimmen im Konzert der Disziplinen. In: Kolesch/Krämer (Hg.), S. 7–15.
- Kolesch, Doris/Krämer, Sybille (Hg.) (2006): Stimme. Annäherung an ein Phänomen. Frankfurt a.M.
- Kotthoff, Helga (1999): Coherent keying in conversational humor: Contextualising joint fictionalisation. In: Bublitz, Wolfram/Lenk, Uta/Ventola, Eija (Hg.): Coherence in spoken and written discourse. How to create it and how to describe it. Amsterdam u.a., S. 125–150.
- Krämer, Sibylle (2005): Nachdenken über die Stimme. In: Jahrbuch Deutsch als Fremdsprache 31, S. 82–91.
- Krämer, Sibylle (2006): Die ‚Rehabilitierung der Stimme‘. Über die Oralität hinaus. In: Kolesch/Krämer (Hg.), S. 269–295.
- Krämer, Sybille (i.d.Bd.): Sprache, Stimme, Schrift: Zur impliziten Bildlichkeit sprachlicher Medien.
- Krech, Eva-Maria/Richter, Günther/Stock, Eberhard/Suttner, Jutta (1991): Sprechwirkung. Grundfragen, Methoden und Ergebnisse ihrer Erforschung. Berlin (Ost).
- Laver, John (1980): Phonetic evaluation of voice quality. San Diego, S. 37–47.
- Laver, John (2000): The phonetic description of voice quality. Cambridge.
- McAllister-Viel, Tara (2006): Toward an intercultural/interdisciplinary approach to training actor's voices. Thesis for the degree of PhD-Performance Practice. Unveröff. Typoskript. Univ. Exeter.
- Mersch, Dieter (2006): Präsenz und Ethizität der Stimme. In: Kolesch/Krämer (Hg.), S. 211–236.
- Meyer-Kalkus, Reinhart (2001): Stimme und Sprechkünste im 20. Jahrhundert. Berlin.
- Meyer-Kalkus, Reinhart (2004): Literatur für Stimme und Ohr. In: Felderer (Hg.), S. 173–186.
- Monrad-Krohn, Georg Herman (1947). Dysprosody or altered ‚melody‘ of language. In: Brain 70, S. 405–415.
- Müller, Ursula (2002): Russisch. In: Hirschfeld, Ursula/Müller, Ursula/Kelz, Heinrich P. (Hg.): Phonetik international. Von Afrikaans bis Zulu. Kontrastive Studien für Deutsch als Fremdsprache. Internet: www.phonetik-international.de/p-phonetik (Stand: August 2009).

- Nawka, Tadeus/Anders, Lutz Christian (1996): Die auditive Bewertung heiserer Stimmen nach dem RBH-System. Doppel-Audio-CD mit Stimmbeispielen. Stuttgart/New York.
- Nawka, Tadeus/Evans, Ruth (2005): RBH-Training und Diagnostik. Forchheim (CD-ROM)
- Nebert, Augustin Ulrich (2007): Tonhöhe und Sprechstimme – Unterschiede in Mutter- und Fremdsprache. In: Zeitschrift für Interkulturellen Fremdsprachenunterricht 12, 2, Internet: <http://zif.spz.tu-darmstadt.de/jg-12-2/navigation/startbei.htm> (Stand: August 2009).
- Nebert, Augustin Ulrich (2009): Stimmübungen im Fremdsprachenunterricht? In: Deutsch als Fremdsprache. Heft II, S. 105–113.
- Neppert, Joachim (1999): Elemente einer Akustischen Phonetik. Hamburg.
- Neuber, Baldur (2002): Prosodische Formen in Funktion. (= Halesche Schriften zur Sprechwissenschaft und Phonetik 7). Frankfurt a.M. u.a.
- Nothdurft, Werner (2004): ‚Wahrnehmung‘ als Denkmuster für Verstehen in der Sprachtheorie Karl Bühlers. In: Ehlich/Meng (Hg.), S. 125–136.
- Paeschke, Astrid (2003): Prosodische Analyse emotionaler Sprechweise. Berlin.
- Peters, John Durham (2004): The Voice and modern media. In: Kolesch, Doris/Schödl, Jenny (Hg.): Kunst-Stimmen. (= Theater der Zeit: Recherchen 21). Berlin, S. 85–100.
- Pfau, Eva-Maria/Streubel, Hans-Gerhard (Hg.) (1982): Die Behandlung der gestörten Sprechstimme – Stimmfunktionstherapie. Leipzig.
- Poe, Edgar Allan (1841 [2009]): Die Morde in der Rue Morgue (Originaltitel: The murders in the Rue Morgue). [In: Detektivgeschichten. Aus dem Englischen von Sophie Zeitz. 2009. München, S. 7–59].
- Redecker, Beate (2008): Persuasion und Prosodie. (= Halesche Schriften zur Sprechwissenschaft und Phonetik 25). Frankfurt a.M.
- Rothe, Friederike (2006): Zwischenmenschliche Kommunikation: Eine interdisziplinäre Grundlegung. Wiesbaden.
- Schäfer, Anke (2003): Interaktive Funktionen des Sprechausdrucks am Beispiel eines Peergroup-Gesprächs. Unveröff. Diplomarb. Univ. Halle a.d.S.
- Scherer, Klaus R. (1995): Expression of emotion in voice and music. In: Journal of Voice 9, S. 235–248.
- Scherer, Klaus R. (2003): Vocal communication of emotion: a review of research paradigms. In: Speech Communication 40, S. 227–256.
- Scherer, Klaus R./Banse, Rainer/Wallbott, Harald G. (2001): Emotion inferences from vocal expression correlate across languages and cultures. In: Journal of Cross-Cultural Psychology 32 (1), S. 76–92.
- Schmitt, Reinhold/Deppermann, Arnulf (i.d.Bd.): Die multimodale Konstitution eines imaginären Raums als interaktive Problemlösung.

- Schneider, Berit/Bigenzahn, Wolfgang (2007): *Stimmdiagnostik. Ein Leitfaden für die Praxis.* Wien/New York.
- Schwarze, Cordula (2009): *Formen und Funktionen von Topoi im Gespräch.* Phil. Diss. (unveröff. Typoskript). Univ. Halle a.d.S.
- Seidner, Wolfram/Wendler, Jürgen (1982): *Die Sängerstimme. Phoniatische Grundlagen für die Gesangsausbildung.* Berlin (Ost).
- Selting, Margret (1995): *Prosodie im Gespräch: Aspekte einer interaktionalen Phonetik der Konversation.* (= Linguistische Arbeiten 329). Tübingen.
- Selting, Margret (1997): *Interaktionale Stilkunde: Methodologische Aspekte der Analyse von Sprechstilen.* In: Selting, Margret/Sandig, Barbara (Hg.): *Sprech- und Gesprächsstile.* Berlin/New York, S. 9–44.
- Sonninen, Aatto (1970): *Phoniatric viewpoint on hoarseness.* In: *Acta Otolaryngologica* 263, S. 68–81.
- Stock, Eberhard (1987): *Ergebnisse der Wirkungsforschung im Bereich von Stimme und Artikulation.* In: Krech, Eva-Maria/Suttner, Jutta/Stock, Eberhard (Hg.): *Ergebnisse der Sprechwirkungsforschung.* (= Kongress- und Tagungsberichte der Martin-Luther-Universität Halle-Wittenberg 19/F 67). Halle a.d.S., S. 50–124.
- Stock, Eberhard (1991): *Emotionserkennung und Stimme.* In: Kutter, Uta/Wagner, Roland (Hg.): *Stimme.* (= Sprache und Sprechen 25). Frankfurt a.M., S. 173–181.
- Stock, Eberhard (in Zusammenarbeit mit J. Suttner) (1991): *Wirkungen des Stimm- und Sprechausdrucks.* In: Krech/Richter/Stock/Suttner, S. 59–142.
- Stock, Eberhard (1993): *Zur Rolle der ‚lautlichen Qualität‘ in der sprechsprachlichen Kommunikation.* In: Richter, Günther (Hg.): *Methodische Grundfragen der Erforschung gesprochener Sprache.* Frankfurt a.M., S. 77–85.
- Stock, Eberhard (1996): *Deutsche Intonation.* Berlin u.a.
- Stock, Eberhard/Veličkova, Ludmila (2002): *Sprechrhythmus im Russischen und Deutschen.* Frankfurt a.M. u.a.
- Terhardt, Ernst (1998): *Akustische Kommunikation: Grundlagen mit Hörbeispielen.* Berlin u.a.
- Tischer, Bernd (1993): *Die vokale Kommunikation von Gefühlen.* Weinheim.
- Traill, Anthony (1986): *The laryngeal sphincter as a phonatory mechanism in !xóo~bushman.* In: Singer, Ronald/Lundy, John K. (Hg.): *Variation, culture and evolution in African populations. Papers in honour of Dr Hertha de Villiers.* Johannesburg, S. 123–131.
- Trojan, Felix (1948): *Der Ausdruck von Stimme und Sprache. Eine phonetische Lautstilkunde.* Wien.
- Trojan, Felix (1975): *Biophonetik.* Mit einem Beitrag von Günter Tembrock; hrsg. u. vervollständigt von Herbert Schendl. Mannheim/Wien/Zürich.
- Trubetzkoy, Nikolaj S. (1939 [1989]): *Grundzüge der Phonologie.* (= Travaux du Cercle Linguistique de Prague 7). Prag [7. Aufl. 1989. Göttingen].

- Vieregge, Wilhelm H. (1989): *Phonetische Transkription. Theorie und Praxis der Symbolphonetik.* (= Zeitschrift für Dialektologie und Linguistik – Beihefte 60). Stuttgart.
- Vieregge, Wilhelm H. (1996): *Patho-Symbolphonetik: auditive Deskription pathologischer Sprache.* Unter Mitarbeit von Johannes Pahn und Harm K. Schutte. (= Zeitschrift für Dialektologie und Linguistik – Beihefte 100). Stuttgart.
- Waldenfels, Bernhard (2006): *Das Lautwerden der Stimme.* In: Kolesch/Krämer (Hg.), S. 191–210.
- Wendler, Jürgen (1996): *Geleitwort.* In: Nawka, Tadeus/Anders, Lutz Christian: *Die auditive Bewertung heiserer Stimmen nach dem RBH-System. Doppel-Audio-CD mit Stimbeispielen.* Stuttgart/New York, S. 3–4.
- Wendler, Jürgen/Rauhut, Andreas/Krüger, Horst (1986): *Classification of voice qualities.* In: *Journal of Phonetics* 14, S. 483–488.
- Wendler, Jürgen/Seidner, Wolfram/Kittel, Gerhard/Eysoldt, Ulrich (1996): *Lehrbuch der Phoniatrie und Pädaudiologie.* 3., völlig neu bearb. u. erw. Aufl. Stuttgart/New York.
- Wendt, Beate (2007): *Analysen emotionaler Prosodie.* (= Hallesche Schriften zur Sprachwissenschaft und Phonetik 20). Frankfurt a.M. u.a.
- Wendt, Beate/Bose, Ines/Sailer, Michael/Scheich, Henning/Ackermann, Herrmann (2007): *Speech rhythm of a woman with Foreign Accent Syndrome (FAS).* In: Barry, William J./Trouvain, Jürgen (Hg.): *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS) Saarbrücken, Aug. 6–10, 2007.* Saarbrücken. Internet: www.icphs2007.de (Stand: August 2009).
- Werner, Heinz (1932): *Grundfragen der Sprachphysiognomik.* Leipzig.
- Westphal, Kristin (2002): *Wirklichkeiten von Stimmen: Grundlegung einer Theorie der medialen Erfahrung.* Frankfurt a.M. u.a.
- Winkler, Christian (1969): *Deutsche Sprechkunde und Sprecherziehung.* 2., überarb. u. erw. Aufl. Düsseldorf.
- Zilliken, Franziska (1991): *Beziehungskonstituierende Wirkungen des Sprechausdrucks.* Frankfurt a.M. u.a.