

Die psychologische Wirkung von Stimme und Sprechweise

Geschlecht, Alter, Persönlichkeit, Emotion und audiovisuelle Interaktion

Von *Walter Sendlmeier*

Einleitung

Mündliche Sprache ist trotz der Verfügbarkeit aller modernen Medien – einschließlich des Internet – immer noch die mit Abstand im Alltag am häufigsten genutzte Form der menschlichen Kommunikation; und sie ist die natürlichste Form der Kommunikation. Aus dieser ursprünglichen Form heraus entstand, entwicklungsgeschichtlich betrachtet, erst sehr viel später die Schriftlichkeit als sekundär abgeleitetes Medium.

Die Stimme eines Menschen ist ganz individuell. Es dürfte tatsächlich keine zwei Menschen mit derselben Stimme geben. Diesen Tatbestand macht man sich schon seit einigen Jahrzehnten im Rahmen kriminaltechnischer Analysen bei der Sprechererkennung zunutze. Neben der bloßen Bestimmung der Identität eines Menschen anhand der Stimme lassen sich viele weitere Eigenschaften über einen Menschen anhand der Stimme und der spezifischen Sprechweise ableiten. Das Alter, das Geschlecht, der Bildungsgrad, die regionale und die soziale Herkunft, der gesundheitliche sowie der momentane emotionale Zustand können sehr zuverlässig aus dem Stimmklang und der

Sprechweise eines Menschen entnommen werden. So erkennen wir sofort an der Sprechweise eines Menschen, ob er gerade eher freudig, traurig, ängstlich oder ärgerlich ist – und zwar auch dann, wenn der Sprecher uns über den Inhalt etwas anderes vormachen möchte. Mit dem Inhalt von Wörtern können Menschen sehr leicht lügen; Gefühlszustände und Charaktermerkmale sind sehr viel schwieriger im stimmlichen und sprecherischen Ausdruck zu verstellen. Nur sehr gute Schauspieler sind in der Lage, Stimmungen und Gefühle auch im Stimmklang und Sprechausdruck auf eine vollständig glaubwürdige Art darzustellen.

Neben den oben genannten Eigenschaften von Menschen, die uns über die Stimme und die Sprechweise zugänglich sind, ist die Stimme eines Sprechers für den Hörer immer auch Ausdruck seiner Persönlichkeit im Sinne überdauernder charakterologischer Eigenschaften. Das Wort *Person* ist aus dem Griechischen *per sona* abgeleitet, was *durch den Klang* bedeutet. Gemeint ist der Klang der Stimme. Und in der Tat liefert uns die Stimme eines Menschen viel mehr Hinweise auf seine Persönlichkeitsstruktur als die visuelle Erscheinung. Selbst die Augen, die oft als Spiegel der

Seele betrachtet werden, können nicht so subtil und variantenreich die Persönlichkeit eines Menschen zum Ausdruck bringen. Die Dominanz der visuellen Medien hat in den letzten Jahrzehnten dazu beigetragen, dass der Mensch nahezu auf ein »Augentier« reduziert und oberflächliche visuelle Informationen in den Vordergrund gerückt wurden. In der Alltagskommunikation sollten jedoch alle Informationen genutzt werden, die Hinweise auf die Innerlichkeit eines Menschen liefern; und diese Hinweise sind zahlreich und weitgehend unverfälscht in der Stimme und Sprechweise eines Menschen enthalten.

Ob der Sprechbewegungsablauf in seinem rhythmischen Erscheinungsbild eher unruhig und labil oder eher gleichmäßig und fließend erscheint, kann von Hörern mit großer Leichtigkeit eingeschätzt werden. Solche Einschätzungen sind es letztlich, die in das unbewusst gebildete Gesamturteil einfließen. Hinzu kommt, dass es den meisten Menschen leichter fällt, Gefühle oder Einstellungen durch die Stimme oder in der Sprechweise auszudrücken, als sie in Worte zu fassen. Dies ist dadurch begründet, dass diese *nonverbale* Art der Kommunikation die ursprünglichere ist. Schon bevor der Mensch eine voll entwickelte Sprachfähigkeit besaß, konnte er durch Affektlaute und Warnschreie die für das Überleben notwendigen Formen der Kommunikation realisieren.

Geschlechtsspezifische Wirkungsweisen

Will man allgemeine Hinweise zur Wirkung von Stimme und Sprechweise geben, so muss zumindest nach geschlechtsspezifischen Eigenarten unterschieden werden. Bei allen individuellen Verschiedenheiten einzelner Sprecher/-innen lassen

sich zwischen den beiden Geschlechtern auch einige systematische Unterschiede aufzeigen. Zudem können bestimmte Stimm- und Sprechmerkmale ganz unterschiedliche Wirkungen auslösen, je nachdem, ob sie von einem Mann oder einer Frau erzeugt werden.

An der Stimme können wir in aller Regel das Geschlecht sehr zuverlässig erkennen. Eine wichtige Ursache dafür sind die unterschiedlichen biologischen Prädispositionen: Längere und dickere Stimmlippen führen dazu, dass Männer mit einer tieferen mittleren Stimmlage sprechen als Frauen (120 Hz gegenüber 220 Hz). Doch die mittlere Grundfrequenz, mit der ein/e Sprecher/-in spricht, kann bis zu einem gewissen Grad selbst gewählt werden. Dies geschieht in der Regel unbewusst. Durch Spannung bzw. Entspannung der an der Stimmgebung beteiligten Muskeln im Kehlkopf kann die Grundfrequenz innerhalb der jeweiligen biologischen Gegebenheiten erheblich variieren. Mit Blick auf professionelle Sprecher/-innen im Rundfunk und Fernsehen hat sich in den 1980er Jahren die Sichtweise verbreitet, dass tiefe Stimmen generell positiver wirken, da in einigen Untersuchungen höhere Frauenstimmen nachteilig beurteilt worden waren. Eine tiefe Stimme gilt als angenehm, kompetent und vertrauenswürdig. In der Tat ist in den vergangenen Jahrzehnten zu beobachten, dass Frauenstimmen generell tiefer geworden sind. Es spielen offensichtlich soziokulturelle Gründe und Vorurteile eine nicht unerhebliche Rolle. Dieses Phänomen hat sogar dazu geführt, dass Nachrichtensprecherinnen, sobald die Mikrophone eingeschaltet sind, ihre Sprechstimmlage von ihrer konstitutionell gegebenen hin zu einer tieferen, männlicheren ändern. Inwieweit generell tiefere Frauenstimmen positiver auf Hörer wirken, ist damit jedoch

nicht beantwortet. Neuere Untersuchungen deuten darauf hin, dass auch eine zu tiefe Sprechstimmlage bei Frauen zu einer negativen Bewertung führt, da sie möglicherweise zu »unweiblich« klingt.¹

Interessanterweise haben sich trotz aller Veränderungen des Frauenbildes im westeuropäischen Kulturraum viele Geschlechtsstereotype erhalten. Die Wiederholung einer Untersuchung der Sozialpsychologen Hofstätter und Lübbert aus dem Jahre 1958² in 2008, in dem anhand eines Polaritätenprofils die Attribute *typisch männlich* versus *typisch weiblich* beurteilt wurden, ergab, dass 23 von 24 erfragten Eigenschaften bezogen auf Männer und Frauen immer noch signifikant unterschiedlich eingestuft wurden.³ Die Geschlechterwahrnehmung wird also auch heutzutage stark von Stereotypen bestimmt; und für die Beurteilung von Männer- und Frauenstimmen gelten unterschiedliche Maßstäbe. Zu tiefe Frauenstimmen werden von vielen Männern als zu männlich eingestuft; gleichzeitig erscheint vielen Männern bei Frauen aber auch die sehr hohe »Klein-Mädchen-Stimme« als nicht mehr zeitgemäß. In anderen Kulturen – wie z.B. im Japanischen – gilt diese sehr hohe Stimmlage noch als Norm für Frauenstimmen. Sie wird als schwach, unsicher, hilfsbedürftig und subdominant eingestuft. Sicherlich gibt es auch im hiesigen Kulturkreis Männer, die eine solche Stimmlage positiv bewerten. Wie diese Männer auch immer im Einzelnen beurteilen mögen, wichtig ist, dass Frauen sich der unterschiedlichen Stimmlagen und der damit verknüpften Wirkungen bewusst werden. Um effektiv zu kommunizieren und um Missverständnisse zu vermeiden, sollte man sich selbst Klarheit verschaffen, welche Signale man aussendet. Mit einer mittleren Sprechstimmlage von ca.

170 bis 220 Hz braucht sich eine Frau keine weiteren Gedanken um ihre Stimmlage zu machen; dies ist eine gute Ausgangslage für zahlreiche Situations- und Adressaten spezifische Variationen.⁴

Generell gilt, dass Männer wie Frauen versuchen sollten, in ihrer Indifferenzlage zu sprechen – einer konstitutionell mittleren Stimmlage, in der sie ohne viel Anspannung und Anstrengung auch längere Zeit sprechen können. Durch Aufregung oder Angst entsteht jedoch oft eine Erhöhung des Muskeltonus, der die Stimme nach oben rutschen lässt. Oft wird in solchen Situationen nicht mehr richtig ausgeatmet und die Spannung nimmt noch weiter zu. Es ist davon auszugehen, dass die hörbare Anspannung, die durch die Anspannung der an der Stimmerzeugung beteiligten Muskeln bedingt ist, vom Zuhörenden körperlich mitempfunden wird. Wenn also ein/e Sprecher/-in über einen längeren Zeitraum mit angespannter Stimme spricht, kann dies auch für Hörer/-innen eine physische Anstrengung bedeuten, da die Anspannung des Sprechenden auf die Zuhörenden übertragen wird. Die Folge ist eine negative Wirkung der gesamten Sprecherpersönlichkeit.

Eine zu tiefe Stimmlage für Männer gibt es eigentlich nicht. Tiefe Männerstimmen werden in aller Regel positiv bewertet. Und als entspannt, gütig, Vertrauen erweckend, glaubwürdig, kompetent und Autorität ausstrahlend interpretiert. Liegen Männerstimmen dagegen deutlich über dem Mittelwert männlicher Sprecher, so stellt sich eine negative Wirkung ein. Sie werden dann als überspannt, unglaubwürdig, unsicher, Aggression auslösend und erregt empfunden.

Neben systematischen Unterschieden in der mittleren Stimmlage lassen sich auch geschlechtsspezifische Unterschiede in

der Art und dem Ausmaß der Tonhöhenvariation beobachten. Männer weisen in aller Regel eine monotonere Intonation auf. Frauen zeichnen sich in ihrer Sprechweise durch mehr Tonhöhenvariation und mehr Dynamik aus – sie sprechen also melodischer als Männer, was von Hörern als emotional expressiver wahrgenommen wird. Frauen werden als verspielter, abwechslungsreicher und emotionaler, Männer als ernster, eintöniger und rationaler wahrgenommen.

Auch auf der segmentalen Ebene sind Unterschiede in der Aussprachegenauigkeit festzustellen. Frauen artikulieren mit einer größeren Aussprachegenauigkeit als Männer. Vor allem junge Männer scheinen oft ihre Nachlässigkeiten in der Aussprache zu kultivieren; »Nuscheln« wird von manchen Männern als Ausdruck von »Coolness« gesehen – frei nach dem Motto: »Schaut her, ich hab' es nicht nötig, mich anzustrengen«. Die gewünschte Wirkung auf das andere Geschlecht wird jedoch meist nicht erfüllt.

Ein sehr markantes Merkmal, das häufig bei Frauen anzutreffen ist und eher negativ wirkt, ist das Nach-oben-Gehen mit der Stimme am Satzende, obwohl es sich nicht um eine Frage handelt. Wird ein Aussagesatz – eine Feststellung – zum Satzende hin mit ansteigender Stimme, also mit Frageintonation gesprochen, so wirkt dies unsicher. Von Hörern wird zum Ende einer Feststellung oder Mitteilung in Aussageform ein finales Absinken der Stimme erwartet. Häuft sich eine zum Äußerungsende hin ansteigende Art der Satzmelodie, so entsteht beim Hörer der Eindruck, die Sprecherin bitte ängstlich darum, ihre Äußerungen zu akzeptieren und zu bejahen. Wird auf der letzten Silbe einer Äußerung die Stimme wieder nach oben geführt, so wird dies häufig durch

eine Dehnung der letzten Silbe begleitet, was zusätzlich Unsicherheit signalisiert. Zudem fehlt dadurch für Hörer eine adäquate Gliederung in Sinnabschnitte. Wer sicher und selbstbewusst wirken möchte, sollte ein Ansteigen der Satzmelodie zum Satzende hin unbedingt vermeiden – es sei denn, es handelt sich um eine Frage.

Eine andere Art des fehlenden finalen Absinkens ist im Kontext öffentlicher Diskussionen anzutreffen: Dort wird das Obenbleiben in der Satzmelodie nicht selten bewusst eingesetzt – etwa in Interviews oder politischen Diskussionen –, um mit einer solchen progredienten Stimmführung zu verhindern, dass Gesprächspartner in den eigenen Redefluss vorzeitig einfallen und so einen Sprecherwechsel erzwingen.

Auch das Ausmaß und die Wirkung einer behauchten Stimme sind für Männer und Frauen sehr unterschiedlich. Bei einer behauchten Stimme schließen die Stimmlippen nicht vollständig; es entsteht ein glottaler Spalt; die Stimmlippen vibrieren dann ineffizient. Sehr viel Luft wird verbraucht, und in der Glottis, dem Raum zwischen den Stimmlippen, entstehen zusätzlich zum periodischen Stimmmantel Turbulenzen, die sich als leises Rauschen, eben die hörbare Behauchung, niederschlagen. Dieser Effekt tritt besonders stark bei unbetonten Silben und am Ende von Äußerungen auf. Frauenstimmen werden gegenüber Männerstimmen öfter als behaucht eingestuft. Interessanterweise werden behauchte Frauenstimmen als erotisch wahrgenommen, wobei die Eigenschaft *erotisch* mit den Beurteilungen der Stimme als sympathisch, angenehm und mit einem positiven Gesamteindruck hoch korreliert. Bei Männern wird eine behauchte Stimme hingegen als unsympathisch und im Gesamteindruck negativ bewertet.⁵

Durch Hinzufügen der Eigenschaft *Nasalität* kann die erotische Wirkung noch verstärkt werden. Nasalität entsteht, wenn über die Nasalkonsonanten /m/ oder /n/ hinaus auch bei der Bildung anderer Laute, insbesondere der Vokale, das Gaumensegel abgesenkt wird und so der Nasenraum als zusätzlicher Resonanzraum den Klang beeinflusst, wie etwa bei den französischen Nasalvokalen. Der typische Schnupfenklang, bei dem die Nasalkonsonanten gerade nicht durch die Nase artikuliert werden können, hat mit dieser Nasalität nichts gemeinsam. Durch das abgesenkte Gaumensegel verändert sich der Stimmklang hin zum Ausdruck sinnlicher Erregung, wie z.B. den des Nahrungsgenusses, des körperlichen Wohlbehagens und auch der sexuellen Sinnlichkeit. Gerade in Kombination mit Behauchung tritt dieser Effekt deutlich in Erscheinung. Nasalität allein, ohne Behauchung, kann aber auch – speziell bei männlichen Sprechern – den Eindruck von Arroganz hervorrufen.

Unterschiede im Beurteilungsverhalten zwischen Hörern und Hörerinnen sind im Allgemeinen kaum zu finden. Mehrfach konnte in Untersuchungen bestätigt werden, dass sich Hörmuster über soziale Kategorien hinweg recht homogen verhalten. Solche Hörmuster sind soziokultureller Natur und haben sich im Laufe der Geschichte in einer Sprachgemeinschaft herausgebildet. Neben der Persönlichkeit eines Sprechers/einer Sprecherin und den Hörmustern ist für den Sprechstil auch die Sprechsituation von großer Bedeutung. So unterscheidet sich der Sprechstil in der Regel danach, ob man mit einem Kind oder Erwachsenen, privat oder öffentlich, mit einer vertrauten oder fremden Person spricht. Ein Sprechender nimmt so genannte *Sprechrollen* ein, die mit sozialen Rollen verbunden sind. Da Frauen

und Männer von klein auf soziokulturell bedingt unterschiedlich behandelt werden und grundverschiedene Erwartungen und Anforderungen an sie gestellt werden, ist davon auszugehen, dass es geschlechtsspezifische Sprechrollen gibt. Durch die existierenden Hörmuster wird erwartet, dass diese auch eingenommen werden; Sprecher/-innen werden – ob sie es wollen oder nicht – von Hörern beiderlei Geschlechts nach unterschiedlichen Kriterien beurteilt.

Alter und Stimme

Mit zunehmendem Alter kommt es in den Kehlkopfknorpeln zu Verknöcherungen (Ossifikationen). Dies wirkt sich nachteilig auf die Geschmeidigkeit der Stellknorpelbewegungen und damit auf die Position der Stimmlippen aus. Zudem kommt es zu einer Erschlaffung der Bänder der Verbindungskapseln. Beide Aspekte zusammen haben eine Änderung der Öffnungsgröße der Glottis zur Folge, die sich als glottale Fehlfunktion bemerkbar macht. Bei beiden Geschlechtern treten ähnliche physiologische Veränderungen im alternden Kehlkopf auf, jedoch mit verschiedenen zeitlichen Abläufen. Bei Männern treten bereits in der dritten Dekade erste Veränderungen im Gewebe auf, ab der fünften oder sechsten Dekade sind diese dann deutlich ausgeprägt. Frauen hingegen weisen erst ab der fünften Dekade Veränderungen auf, die außerdem insgesamt schwächer ausgebildet sind.⁶

Die obere Zellschicht der Stimmlippen, das Epithel, flacht mit zunehmendem Alter ab. Ab dem 40. Lebensjahr reduziert sich das Gewebe der mittleren Schicht. Die größten Veränderungen vollziehen sich ab dem 50. Lebensjahr in der tiefen Schicht, wo eine Verdickung kollagener Fasern den Verlust

einer linearen Form hin zu einem zerfaserten Gebilde hervorruft. Diese Veränderungen in den tieferen Schichten sind im Alterungsprozess von Frauen nicht vorhanden. Der für die Feinregulierung der Phonation zuständige Vokalismuskel verliert an Kraft. Dies führt zusammen mit der Umwandlung des Gewebes zu einem Verlust an Spannung, Elastizität und Masse der Stimmlippen und kann Unregelmäßigkeiten im Schwingungsverhalten nach sich ziehen, was wiederum perceptiv zu einem rauen Klangeindruck führt. Ebenso können Lücken im Glottisverschluss entstehen, die zu einer rauschartigen Behauchung führen.

Dünnere und steifere Stimmlippen, wie sie sich durch die veränderte Struktur des Bindegewebes ergeben, schwingen schneller und mit kleinerer Amplitude. Dies entspricht der Tatsache, dass ältere Männer höhere Grundfrequenzen aufweisen. Bei Frauen hingegen sinkt die Sprechstimmlage im Alter. Sowohl für junge als auch für ältere Frauen ist häufig das Auftreten von unvollständigen Glottisverschlüssen beobachtet worden. Hier zeigten sich aber Unterschiede in der Art der unvollständigen Glottisverschlüsse, die nicht immer über die gesamte Länge entstehen. Bei jungen Frauen überwiegt der so genannte *posterior chink*, ein hinterer Spalt, und bei älteren Frauen der so genannte *anterior gap*, eine vordere Lücke. Eine befriedigende Erklärung für diese Unterschiede wurde bislang nicht gefunden.⁷ Die unterschiedliche Art der Behauchung kann aber auditiv unterschieden werden.

Zu berücksichtigen ist in diesem Zusammenhang jedoch auch, dass altersbedingte hormonelle Veränderungen Auswirkungen auf den Stimmklang haben. So ist nicht nur in der Pubertät mit dem Wachstum des Kehlkopfes ein wesentlicher Einfluss gegeben, sondern auch die

hormonellen Veränderungen insbesondere bei Frauen aufgrund des Klimakteriums bzw. nach der Menopause stellen eine wichtige Ursache für stimmliche Veränderungen dar. Insgesamt ist zu beachten, dass Alterungserscheinungen nicht nur vom kalendarischen Alter abhängig sind, sondern in starkem Maße auch vom gesundheitlichen Status und dem Gesamtzustand des Organismus. So gibt es denn auch mehr alte Menschen als ausgeprägt alt klingende Stimmen.

Die Stimme älterer Sprecherinnen ist zittriger, tiefer, rauer und behauchter. Beim Vorlesen machen ältere Sprecher/-innen mehr Sprechpausen, sprechen insgesamt langsamer und artikulieren weniger genau.⁸ Unabhängig vom Geschlecht verliert die Stimme im Alter in erheblichem Maße an Stabilität. Das lautere Sprechen der älteren Sprecher/-innen wird möglicherweise als Kompensation für die schwache und behauchte Stimme eingesetzt. Es könnte aber auch eine Folge der Schwerhörigkeit älterer Menschen sein. Der mittlere Öffnungsquotient, der anhand elektrolotographischer Messungen ermittelt wird, ist unabhängig vom Geschlecht höher bei älteren Sprechern und kann als Hinweis auf fortschreitende laryngale Degeneration im Alter interpretiert werden.⁹ Die Analyse der lokalen Variationen des Sprechtempos in Spontansprache zeigt, dass besonders die schnellen Abschnitte verlangsamt werden, was zu einer Einschränkung der Variationsbreite und somit zu einer gleichförmigeren Sprechweise führt.¹⁰

Die sich ergebende Form der Glottispulse ist eher sinusförmig als sägezahnförmig, was spektral mit einer stärkeren bis vollständigen Dämpfung der höheren Harmonischen einhergeht. Diese Art der Anregung tritt bei alten Menschen häufiger und ausgeprägter auf. Im Gegensatz

dazu entsteht bei jüngeren Sprechern durch stärkere Rückstellkräfte eine impulsförmigere Anregung, die eine geringere Dämpfung hin zu den höheren Harmonischen aufweist. Für Hörer ist dieser Unterschied – unabhängig von der Tonhöhe des Grundtons – in Form eines dumpferen (älteren) oder aber eines helleren (jüngeren) Stimmklangs erfahrbar. Die Beschreibung der physiologischen Korrelate legt diese altersabhängigen Veränderungen im Grunde genommen nahe.

Emotionaler Ausdruck durch Stimme und Sprechweise

Freude, Trauer, Ärger (Wut) und Angst sind vier Emotionen, die in der einschlägigen psychologischen Forschung unstrittig als Basisemotionen eingestuft werden. In vielen Ansätzen wird noch der Zustand *Langeweile* hinzugefügt. Diese fünf Basisemotionen sind Gegenstand der folgenden Betrachtungen. Aus unserer Alltagserfahrung wissen wir, dass der emotionale Zustand von Sprechern auch unabhängig vom Inhalt des Gesagten erkannt werden kann. Was ändert sich nun an unserer Sprechweise, wenn wir ärgerlich, freudig, traurig oder ängstlich sprechen?

Zunächst zum stimmlichen Ausdruck: Beim Schwingen der Stimmlippen verläuft das Schließen der Glottis wesentlich schneller als ihr Öffnen, was neben den oben beschriebenen muskulären Rückstellkräften auch an der durch den Bernoulli-Effekt bewirkten Sogwirkung liegt. Die Analysen emotionaler Sprechweisen zeigen, dass sich je nach Emotion die Form der Glottispulse verändert, die wiederum zu einer Veränderung des Glottisspektrums führt. Diese aber ist dann für Hörer als bestimmter emotionaler Stimmklang zu identifizieren.¹¹

Die Glottispulse zeigen bei den Emotionen Ärger und Freude steilere Flanken und längere geschlossene Phasen als die entsprechenden Anregungssignale bei neutraler Sprechweise. Bei Ärger sind die geschlossenen Phasen am längsten. Die Anregung ist also impulsförmiger als bei neutralen Äußerungen. Diese impulsförmigere Anregung wiederum führt im spektralen Bereich (also in der Energieverteilung auf die Obertöne) dazu, dass hier im mittleren und hohen Frequenzbereich mehr Energie vorhanden ist als bei der neutralen Sprechweise. Anders formuliert: Die spektrale Dämpfung der höheren Harmonischen ist bei Freude und Ärger geringer als bei neutraler Sprechweise; am geringsten ist sie bei ärgerlicher Sprechweise. Die Grundfrequenz der ärgerlichen und freudigen Äußerungen weist insgesamt eine größere Variationsbreite (F_0 -Range) auf als die neutralen Äußerungen. Die mittlere Stimntonfrequenz wird angehoben – und zwar bei den freudigen Äußerungen noch mehr als bei den ärgerlichen.

Bei trauriger Sprechweise ist die Form der Glottisanregung gerade nicht impulsförmig, sondern sogar fast sinusförmig, bewegt sich also – wenn man die sägezahnähnliche Form der Glottispulse der neutralen Sprechweise als Bezugsgröße nimmt – gerade in die andere Richtung. Im Glottisspektrum führt die sinusförmige Anregung dazu, dass im tieffrequenten Bereich mehr Energie als in den neutralen Äußerungen liegt, während im mittleren und höheren Frequenzbereich deutlich weniger Energie als in den neutralen Äußerungen auftritt. Im Anregungssignal sind hochfrequente Rauschanteile enthalten, die Hinweise auf eine Behauchung geben. Eine Ähnlichkeit mit den Eigenschaften älterer Stimmen ist zu beobachten. Zusätzlich treten bei traurigen Sprechern

Laryngalisierungen auf. Die traurigen Äußerungen sind also durch eine knarrende bis behauchte Anregung gekennzeichnet. Eine Grundfrequenzbewegung findet kaum statt; außer bei der Frageintonation ist sie nur abwärts gerichtet. Auffallend ist bei den traurigen Äußerungen auch, dass die Vokale weit weniger Energie als in den neutralen Versionen aufweisen. Die geringe Energie der Vokale und die fast monotone, nur gering fallende Intonation weisen auf eine allgemeine Passivität und geringen Muskeleinsatz hin.

Bei ängstlichen Äußerungen sieht die Form der Anregungsimpulse – soweit überhaupt welche vorhanden sind – jener bei Trauer durchaus ähnlich; die frequenzabhängige Energieverteilung ist bei ängstlichen Äußerungen aber eine andere: Im hochfrequenten Bereich ist hier viel Energie anzutreffen; jedoch handelt es sich hier primär nicht um einen harmonischen Anteil (wie bei Freude und Ärger), sondern um einen turbulenten, d.h. rauschhaften Anteil. Man könnte von einer Art geflüsterter Falsettstimme sprechen. Die mittlere Grundfrequenz liegt recht hoch; sie weist aber nur eine geringe Variationsbreite auf.

Die Form der Glottispulse ähnelt bei gelangweilter Sprechweise der bei normaler (neutraler) Sprechweise sehr. Die Vokale sind wie bei der Trauer sehr energiearm; sie weisen jedoch keine Laryngalisierungen auf und zeigen eine deutlich geringere spektrale Dämpfung als die traurig gesprochenen Äußerungen.

Trauer und Angst sind bei allen Sprechern die Emotionen, die die meisten Verschleifungen und Elisionen – z.T. sogar die Auslassung ganzer Silben – nach sich ziehen. Die Äußerungen mit ärgerlicher Sprechweise zeigen bei allen Sprechern die geringsten Auslassungen und Verschmelzungen. Bemerkenswert ist dies vor allem

deswegen, weil Ärger nach Angst die Emotion ist, die mit der schnellsten Sprechgeschwindigkeit einhergeht. Zahlreiche Untersuchungen belegen, dass ein enger Zusammenhang zwischen Sprechtempo und Reduktionsphänomenen besteht. Dieser Zusammenhang scheint bei ärgerlicher Sprechweise außer Kraft gesetzt zu werden. Fasst man die Beobachtungen zur Artikulationsgenauigkeit zusammen, so ist festzuhalten, dass die Sprechweise bei Angst und Trauer einerseits und Ärger andererseits sehr unterschiedlich ausgeprägt ist. Ärger ist die einzige Emotion, bei der sich im Vergleich zur neutralen Version auch laborierte Elemente in der Sprechweise finden lassen. Ärger geht also häufig mit einer präzisen Artikulation einher. So lassen sich viele Beispiele dafür finden, dass bei Ärger deutlicher artikuliert wird als in neutraler Sprechweise sowie bei allen anderen Emotionen. In diesem Zusammenhang ist auch zu beachten, dass bei ärgerlichen Äußerungen eine Zunahme stark betonter und eine Abnahme unbetonter Silben eintritt. Bei ängstlichen und traurigen Äußerungen tritt genau das Gegenteil ein. Betonte Silben werden in aller Regel mit einer präziseren Artikulation realisiert, unbetonte Silben hingegen mit ausgeprägten Reduktionen.¹²

Bei Trauer und Langeweile ist die hohe Anzahl von konsonantischen und vokalischen Reduktionen besonders bemerkenswert, da die Sprechgeschwindigkeit bei diesen Emotionen geringer ist als bei anderen Emotionen oder der neutralen Sprechweise und daher eine unpräzise Artikulation eigentlich unwahrscheinlich ist. Eine Erklärung für die starke Reduktion bei Trauer und Langeweile ergibt sich aus der geringen Muskelspannung, mit der diese Emotionen aufgrund des geringen Erregungsgrades einhergehen.

Betrachtet man die Satzmelodien emotionaler Äußerungen, so ist anhand der Intonationskonturen zu erkennen, dass eine relativ zuverlässige Diskrimination zwischen Emotionen mit hohem und niedrigem Erregungsgrad möglich ist. Es können deutliche Unterschiede zwischen traurigen, gelangweilten und neutralen Äußerungen einerseits und ängstlichen, freudigen und wütenden Äußerungen andererseits gefunden werden. Darüber hinaus zeigen sich auch signifikante Unterschiede zwischen Langeweile und Trauer sowie zwischen Freude und Wut. Auch Angst ist von anderen Emotionen abgrenzbar. Die Messungen zur Variationsbreite und Stimmlage ergeben, dass ängstliche Äußerungen einen ähnlich schmalen Variationsbereich wie traurige besitzen und die Minima der Grundfrequenz (F_0) zwischen den Akzenten leicht unter den Werten von Freude und Wut liegen. Beide Merkmale zusammen sind ein eindeutiges Indiz für eine ängstliche Äußerung. Zusätzlich dazu weist Angst die kürzesten Akzente, als einzige Emotion eine insgesamt ansteigende F_0 -Kontur und den höchsten finalen F_0 -Wert auf.¹³

Die Analyse ergibt auch signifikante Unterschiede zwischen Trauer und Langeweile. Gelangweilte Äußerungen unterscheiden sich von anderen Emotionen durch eine signifikant längere Dauer der Akzente und ein längeres Absinken der Grundfrequenz auf der letzten Silbe. Weiterhin sind die Unterschiede zwischen F_0 -Maxima und F_0 -Minima der Akzente signifikant größer als bei traurigen Äußerungen. Die Grundfrequenzkontur zeigt insgesamt bei Langeweile ein stärkeres Absinken als bei traurigen Sätzen.

Bei der Analyse von Satzmelodien emotionaler Äußerungen können verschiedene Typen von Intonationskonturen innerhalb

einer Emotion gefunden werden. Dies gilt speziell für wütende, freudige und gelangweilte und zu einem geringeren Teil auch für ängstliche Äußerungen. So lässt etwa eine Unterscheidung der Kategorie Ärger in *Wut* und *unterdrückten Ärger* oder die Aufteilung von Trauer in *offene* und *stille Trauer* eine jeweils noch einheitlichere Interpretation zu.

Stimme und Persönlichkeit

Wenn eine Stimme fest und die Satzmelodie lebendig klingen, sind unmittelbar erste Hinweise gegeben, dass Eigenschaften wie Willensstärke, Energie, evtl. auch Durchsetzungsvermögen oder innere Beweglichkeit vorhanden sind. Die Art der Betonung kann dann Aufschluss geben, ob die Energie eher in Richtung des Durchsetzungsvermögens im Sinne eines Befehls tones geht oder in Richtung einer Begeisterungsfähigkeit. Ein minimales Zittern im Stimmklang könnte andeuten, dass es um die Selbstsicherheit dieses Menschen doch nicht so zuverlässig bestellt ist. Obwohl das menschliche Ohr für solche Nuancen hoch sensibel ist, werden unsere Fähigkeiten in diesem Bereich oft stark unterschätzt, weil solche Analysen und Bewertungen im Alltag unbewusst ablaufen und Laien nicht in der Lage sind, ihre Urteile auf einer Metaebene zu begründen. Eine Begründung ist im Alltag auch nicht erforderlich. Wichtiger ist, dass Stimmen unmittelbar auf uns wirken und wir uns solchen Beurteilungen kaum entziehen können. Jeder Mensch kann schon nach einem ersten Eindruck einer Stimme sagen, ob die Stimme nach seinem Gefühl angenehm oder unangenehm klingt.

Fragt man Hörer gezielt nach dem Eindruck eines Stimmklanges und gibt man dazu Beschreibungshilfen in Form unter-

schiedlicher Adjektive, so können Hörer in aller Regel Stimmen schnell und zuverlässig charakterisieren. Auch bei der Wiederholung solcher Befragungen bleiben die Hörer bei ihrem Urteil. Ob eine Stimme eher warm und weich oder aber eher scharf, kalt und hart klingt, kann von Hörern klar eingeschätzt werden.

Da die Stimme einen Einblick in die Seele, in die Innerlichkeit, in den Charakter eines Menschen gibt, ist sie etwas Intimes. Sie verrät evtl. auch unsere Angst, zurückgewiesen zu werden, unsere Sehnsüchte, unsere Aufgeregtheit oder Verlegenheit. Schon durch den ersten Höreindruck von Stimme und Sprechweise erhalten wir ein recht differenziertes »Bild« von einem Menschen. Die Anonymität und Beliebigkeit von personenbezogenen Daten werden durch den spezifischen Stimmklang und die individuelle Art der Satzmelodie und der Aussprache aufgehoben. Der Zugang zu wesentlichen Aspekten der Persönlichkeit ist unmittelbar und weitgehend unverfälscht gegeben. Wer auf einen Menschen mit einer bereits bekannten Stimme zugeht, weiß weit mehr, worauf er/sie sich einlässt als beim Vorliegen einer bloßen schriftlichen Beschreibung oder der Betrachtung eines Bildes. Beides kann sehr leicht manipuliert werden und sagt über das tatsächliche Wesen eines Menschen recht wenig aus.

Die Persönlichkeitsforschung hat sich seit langem um die Messung von stabilen Persönlichkeitsmerkmalen bemüht. Am bekanntesten sind die beiden Merkmale *Extraversion* und *Neurotizismus* geworden. Inzwischen haben viele Untersuchungen immer wieder die Existenz von fünf Dimensionen nachgewiesen, die man auch als Big Five, die großen Fünf, bezeichnet. Ihre Ausprägung bei einzelnen Personen kann sehr zuverlässig durch einen Persönlich-

keitstest – den NEO-PI-R – ermittelt werden. Dieser weltweit in der Forschung und klinischen Praxis am häufigsten eingesetzte Fragebogen zur Messung des Fünf-Faktoren-Modells der Persönlichkeit basiert auf den Ergebnissen jahrzehntelanger faktorenanalytischer Forschung mit umfangreichen Bevölkerungsstichproben und klinischen Probandengruppen. Das revidierte NEO-Persönlichkeitsinventar (NEO-PI-R, nach Costa und McCrae¹⁴) erfasst mit 240 Items die Hauptbereiche interindividueller Persönlichkeitsunterschiede. Der besondere Vorteil dieses Inventars liegt in seiner großen Bandbreite. Es ermöglicht eine umfassende und zugleich detaillierte Persönlichkeitsbeschreibung. Auch für die Erhebung von Fremdbeurteilungen z.B. durch Familienangehörige, nahe stehende Bekannte oder Klinikpersonal steht ein entsprechendes Formular zur Verfügung. In der Praxis wird sehr häufig eine Kurzform dieses Tests eingesetzt, die unter der Bezeichnung NEO-Fünf-Faktoren-Inventar (NEO-FFI) firmiert und ebenfalls sehr zuverlässige Ergebnisse liefert.

Es können folgende fünf Kern-Dimensionen unterschieden werden, die die Basis unserer Persönlichkeit bilden: die Faktoren *Neurotizismus*, *Extraversion*, *Offenheit für Erfahrung*, *Verträglichkeit* und *Gewissenhaftigkeit*. Für die ersten beiden Faktoren, die oft als die wichtigsten eingeschätzt werden, gibt es erste Ergebnisse im Hinblick auf stimmliche und sprechsprachliche Indikatoren.¹⁵

Die Skala Neurotizismus erfasst Unterschiede in der emotionalen Stabilität gegenüber der emotionalen Labilität. Stabilität wird durch Attribute wie stabil, ruhig, sorglos und unempfindlich gekennzeichnet, Labilität dagegen durch Attribute wie angespannt, besorgt, unsicher und sensibel.

Die Skala Extraversion unterscheidet gesellige Menschen, die gesprächig, aktiv und heiter sind und gern auf andere Menschen zugehen von solchen, die eher zurückhaltend sind und oft den Wunsch verspüren, allein zu sein, ohne aber unter sozialer Ängstlichkeit zu leiden.

Sprecher/-innen, die sich in unseren Untersuchungen mithilfe des NEO-FFI selbst als eher ängstlich, unsicher, besorgt und sensibel einschätzen, werden auch von Hörer/-innen aufgrund kurzer Sprachbeispiele dieser Sprecher – ebenfalls mit Hilfe der Frageitems des NEO-FFI – als eher labil eingestuft. Die Sprachbeispiele bestanden aus ca. 15 Sekunden gelesener bzw. spontan gesprochener Rede, deren Inhalt keinerlei Aufschluss über Charaktereigenschaften ergab; in einem weiteren Testteil wurde lediglich der gehaltene Vokal /a/ dargeboten. Je komplexer die sprachlichen Strukturen sind, desto deutlicher treten die Sprechereigenschaften zu Tage. Aber bereits anhand der isoliert erzeugten und wenige Sekunden gehaltenen Vokale konnten signifikante Korrelationen zwischen den Beurteilungen der Persönlichkeitseigenschaften durch die Sprecher selbst und den Fremdbeurteilungen durch Hörer gefunden werden.

Für die Skala Extraversion ergaben sich ebenfalls signifikante Korrelationen zwischen den Selbsteinschätzungen der Sprecher/-innen und den Fremdbeurteilungen durch Hörer, wobei den Hörerurteilen wiederum lediglich die kurzen sprachlichen Äußerungen der Sprecher zugrunde lagen. Die Einstufungen auf den NEO-FFI-Items sind in hohem Maße ähnlich – erneut kam es zu Übereinstimmungen zwischen Selbst- und Fremdbeurteilung

Die Sprechweise der als stabil eingestuften Sprecher/-innen ist gekennzeichnet durch spektrale Eigenschaften, die in hö-

heren Frequenzen mehr Energie und klarere Formantkonturen aufweisen. Dieses Merkmal geht wahrscheinlich auf stärkere Rückstellkräfte der an der Phonation beteiligten Muskeln zurück. Weiterhin ist die mittlere Tonhöhe der als stabil eingestuften Sprecherinnen niedriger als die der als labil beurteilten Sprecherinnen. Die Maße für Stimmstabilität (Shimmer, Jitter und Tremor-Index) sind für die als labil eingestuften Sprecherinnen deutlich höher als bei den als stabil eingestuften.

Die extrovertierten Sprecher weisen eine schnellere und bisweilen deutlichere Artikulation als die introvertierten Sprecher auf. Die Sprechweise der introvertierten Sprecher ist besonders durch eine geringere Dynamik und Grundfrequenzvariation gekennzeichnet; ihre Satzmelodie wirkt insgesamt monotoner. Als extrovertiert wahrgenommene Sprecher/-innen weisen entsprechend eine größere Variationsbreite hinsichtlich prosodischer Merkmale auf. Die extrovertierten Sprecher sprechen auch deutlich lauter als die introvertierten Probanden.

Synchronstimmen

Der Frage nach der stimmlichen Manifestation von Charaktereigenschaften kann man sich auch über eine Analyse von Synchronstimmen nähern. Dadurch kann deutlich werden, wie sehr durch die Stimme und Sprechweise die Interpretation des Charakters einer Filmrolle beeinflusst werden kann.

Am Beispiel des US-amerikanischen Kinofilms *AGAINST THE ROPES* (Die Promoterin; 2004; R: Charles S. Dutton) sind wir der Frage nachgegangen, inwieweit sich der Einsatz einer Synchronstimme verändernd auf den entsprechenden Charakter einer Filmrolle auswirkt bzw. ob Hörer

die verschiedenen Stimmen und Sprechweisen in der Originalversion und in der synchronisierten Fassung unterschiedlich wahrnehmen.¹⁶ Der Film erzählt die Lebensgeschichte einer erfolgreichen, selbstbewussten Frau namens Jackie Kallen (Meg Ryan), die sich als Boxpromoterin in einer Männerwelt zu beweisen versucht. Die deutsche Synchronsprecherin Ryans ist Ulrike Möckel.

Ryan spricht diese Filmrolle mit einer tiefen und rauen Stimme; an Äußerungen kommt es häufig zu einer ausgeprägten Knarrstimme. Die Sprechweise wirkt monoton und ruhig, oft gelangweilt und sogar depressiv. Die Artikulation ist undeutlich und teilweise durch völlige Tilgung der Plosive im Auslaut geprägt. Die deutsche Synchronstimme hat dagegen eine etwas höhere, festere und klarere Stimme. Nur selten ist leichtes Knarren in der Stimme wahrnehmbar. Die Sprechweise wirkt sehr lebendig, zielstrebig und engagiert; die Artikulation ist deutlich und oft mit hartem Stimmeneinsatz. Beide weisen häufige Betonungen auf, allerdings scheinen diese auf unterschiedliche Art erzeugt zu werden.

Ergänzt man die auditive Beschreibung in einem zweiten Schritt durch systematische Hörtests in Form eines aus 20 adjektivistischen Gegensatzpaaren bestehenden Polaritätsprofils zur Erfassung der Sprechwirkung, so ergeben sich folgende Beobachtungen: Die statistischen Auswertungen zeigen deutliche Unterschiede in der Beurteilung der Stimme und Sprechweise der beiden Sprecherinnen. Das US-amerikanische Original wurde von Hörern beider Muttersprachen Deutsch und Englisch insgesamt als monotoner, gleichgültiger, behauchter, weniger energisch, eintöniger, rauher, weniger autoritär, undeutlicher, tiefer, langsamer und schwächer eingestuft als die deutsche Synchronstimme. Aber

sie wurde auch als entspannter, natürlicher und angenehmer empfunden. Demnach erscheint die Umsetzung des Filmcharakters in seiner intendierten Wirkung durch Meg Ryan weniger gelungen als die deutsche Stimme, die der Sprechweise einer starken Geschäftsfrau den Hörerurteilen sowie den in Interviews von Regisseur und Drehbuchautorinnen geäußerten Vorstellungen des Filmcharakters gemäß eher entspricht.

Auch die Ergebnisse der akustischen Analyse weisen deutliche Unterschiede auf. Ryan spricht mit einer für Frauenstimmen sehr tiefen mittleren Grundfrequenz und weist einen relativ geringen Stimmumfang sowie geringe Stimmvariabilität auf. Ihre Intonationsverläufe enthalten viele lange, gerade und flach abfallende Konturlinien. Dies weist auf eine monotone, gleichgültige, wenig energische und eintönige Sprechweise hin. Betonungen werden primär durch Veränderungen der Vokalqualität, Dehnungen von Silben sowie Betonungspausen erzeugt. Sie spricht im Vergleich zu der deutschen Sprecherin langsamer. Die Stimmqualität Ryans ist durch relativ wenig Energie im oberen Spektralbereich, hohe Jitter- und NHR (noise-to-harmonic-ratio)-Werte sowie zahlreiche Laryngalisierungen gekennzeichnet. Dies sind die akustischen Korrelate der auditiv als rau, behaucht und knarrend empfundenen Stimmqualität. Die deutsche Stimme wird insgesamt in einer höheren Tonlage realisiert; zudem weist sie einen größeren Stimmumfang sowie eine größere Standardabweichung auf. Die Intonationsverläufe sind sehr bewegungsreich; ausgeprägt steigende und fallende Intonationsverläufe kennzeichnen vor allem betonte Silben. Dies ist die akustische Grundlage für die Wahrnehmung der Sprechweise durch die Hörer als abwechslungsreich, enga-

giert, energisch, lebendig und hoch. Die Wahrnehmung der Artikulation als sehr deutlich kann durch die dezentralisierte Lage der ersten beiden Formanten bestätigt werden. Betonungen werden primär durch Veränderungen der Intonationskontur auf Silbenebene erzeugt. Aufgrund der höheren Wortlastigkeit der deutschen Sprache muss Ulrike Möckel mehr Text artikulieren als das englische Original, was zu einer höheren Sprechgeschwindigkeit und kürzeren Pausen führt.

Bildhafte Vorstellungen von Radiomoderatoren

Im Rahmen von Untersuchungen zu Stimme und Sprechweise im Hinblick auf Indikatoren für zugrunde liegende oder auch andere wahrnehmbare Eigenschaften von Menschen lässt sich auch die Frage aufwerfen, ob Hörer/-innen sich von Sprechern / Sprecherinnen eine bildhafte Vorstellung machen, wenn ihnen nur Information über den auditiven Kanal zur Verfügung steht. Die sehr wenigen, älteren Arbeiten zu dieser Thematik legen nahe, dass sich Hörer ein mentales Bild von einem Sprecher / einer Sprecherin machen, auch wenn sie diese nicht visuell wahrnehmen können. Neuere Forschung zu dieser Fragestellung liegt nicht vor. Erste eigene Untersuchungen mit drei Sprechern und drei Sprecherinnen zu der Frage, ob sich Radiohörer von Radiomoderatoren eine bildhafte Vorstellung machen, ob sich die mentalen Bilder der Hörer untereinander gleichen und inwieweit die bildhaften Vorstellungen mit bestimmten stimmlichen und sprecherischen Merkmalen in Verbindung gebracht werden können, wurden an unserem Institut an der Technischen Universität Berlin mithilfe eines Phantombildprogramms durchgeführt.¹⁷

Für die im Fragebogen erhobenen körperlichen Merkmale gab es bei den Hörerurteilen deutliche Übereinstimmungen. Die Einschätzungen des Alters, der Körpergröße und selbst der Haarfarbe wiesen große Übereinstimmungen auf. Auch für die Vorstellung der äußerlichen Attraktivität, der Statur (schlank vs. dick) oder des Ausmaßes der eingesetzten Mimik gab es hohe Übereinstimmungen in den Hörerurteilungen.

Die mit dem Phantombildprogramm erhobenen Gesichter wiesen jedoch nur bedingt Ähnlichkeiten auf. Dennoch sind auch hier gewisse Tendenzen zu ähnlichen Annahmen über das Aussehen der Sprecher zu finden. Einige seien hier beispielhaft aufgeführt:

- Sprecher 1, der die tiefste Stimme hat, wurde ein sehr attraktives, maskulines Erscheinungsbild attribuiert. Mit den höchsten Anteilen an bärtigen Gesichtern, sehr markanten Gesichtszügen und eher dunklen Augen scheint er vielen Hörern dem Klischeebild eines attraktiven starken Mannes zu entsprechen. Seine Phantombilder sehen sich am ähnlichsten.

- Für Sprecher 2 ergab sich eine sehr deutliche Tendenz zum Klischeebild des jungen, eher schmalen und modernen Mannes mit aktueller Frisur. Er hat in der Vorstellung der Hörer volle Haare und etwas hellere Augen. Seine Phantombilder ähneln sich ebenfalls in hohem Maße, die Hörer gaben aber an, dass sie das mentale Bild dieses Sprechers nicht so deutlich vor Augen hätten; umso mehr überrascht die große Übereinstimmung.

- Für die Phantombilder von Sprecher 3 sind erhebliche Abweichungen festzustellen, insbesondere hinsichtlich der Breite des Gesichtes. Dennoch wurde dieser Sprecher recht einheitlich als am wenigsten attraktiv eingestuft. Seine – für einen

Mann – sehr hohe Stimmlage und die sehr angespannte Sprechweise könnten die Ursache für das eher negative Gesamtbild dieses Sprechers sein.

– Für die drei Sprecherinnen lassen sich ähnliche Tendenzen aus den Daten ableiten.

Die Korrelationen der Hörerbeurteilungen mit den akustischen Messungen stimmlicher Parameter erbrachten ebenfalls zahlreiche signifikante Zusammenhänge. So führte etwa eine tiefere Stimmlage zu der Beurteilung »attraktiv«; dies galt vor allem für die männlichen Sprecher. Dasselbe gilt auch für eine schnellere Sprechweise. Sprecher mit einem hohen F_0 -Range wurden eher als fülliger eingeschätzt. Gab es in der Satzmelodie eine schnellere Folge von Konturrichtungsänderungen, so wurden die Sprecher als sportlicher, jünger und insgesamt als agiler eingeschätzt. Bei einem hohen Anteil steigender Konturen wurden die Sprecher als weniger sportlich und weniger attraktiv beurteilt. Positiv wurde ein hoher Anteil an fallend-steigenden Konturen wahrgenommen, die stark mit Größe und Attraktivität korrelierten.

Das vegetative Nervensystem

Sind Eigenschaften von Stimme und Sprechweise im Kontext der Übernahme sozialer Rollen primär das Ergebnis gelernter Verhaltensweisen, so sind stimmliche Merkmale als Ausdruck von Persönlichkeitseigenschaften eher eine Mischung aus gelerntem und angeborenem Verhalten. Stimmliche Veränderungen in Abhängigkeit des emotionalen Zustandes sind sicherlich als universell determiniert zu betrachten, wenn auch im Ausprägungsgrad kulturell überformt. Emotionale Zustände wie Trauer oder Ärger werden in manchen Kulturen offener gezeit als in anderen.

Die überdauernden Persönlichkeits- bzw. Charaktereigenschaften und auch die stärker schwankenden emotionalen Zustände stehen nicht nur mit bestimmten Erregungen im zentralen Nervensystem in einem Zusammenhang, sondern werden auch stark durch Aktivitäten der beiden Komponenten des vegetativen Nervensystems – des Sympathikus und des Parasympathikus – beeinflusst. Das vegetative Nervensystem steuert lebenswichtige Funktionen des Körpers wie Kreislauf, Verdauung, Stoffwechsel, Atmung, Sekretion usw. Die Aktivierung von Sympathikus und Parasympathikus führt dabei zu entgegengesetzten Wirkungen. Es ist die abwechselnde Aktivität dieser beiden Mechanismen, die die Regulation der Körperfunktionen ermöglichen. Auch bei Emotionen und bestimmten Charakterausprägungen zeigen sich eher sympathische oder aber parasympathische Aktivitäten.

Es liegen bisher keine Untersuchungen darüber vor, welche direkten Auswirkungen der vegetativen Erregung auf die Physiologie der Sprachproduktion bestehen. Dass es diese aber gibt, ist nahe liegend, wenn man z.B. bedenkt, dass Sprechen ganz wesentlich von der Atmung abhängig ist. Zudem führt sympathische wie parasympathische Erregung zu veränderter Muskelaktivität – besonders der quer gestreiften Muskulatur. Solche Veränderungen des Muskeltonus und der Muskelinnervation betreffen grundsätzlich auch die an Stimmbildung und Artikulation beteiligten Muskeln. Außerdem haben die an der Mimik beteiligten Muskeln, die z.T. mit artikulatorisch bedeutsamen Muskeln identisch sind (Lippen, Wangen, Kiefer), phylogenetisch ihren Ursprung im Atmungsapparat (der Wirbelfische) und weisen direkte Verbindungen zum vegetativen Nervensystem

auf. Es könnten bei den Emotionen, die mit sympathischer Erregung verbunden sind, eine bessere Kontrolle, stärkere Anspannung und größere Beweglichkeit der an der Sprachproduktion beteiligten Organe die Folge sein. Bei übergroßer sympathischer Erregung könnten sich diese Auswirkungen jedoch auch derart steigern, dass sowohl im laryngalen als auch im supralaryngalen Bereich extreme Anspannung und Engbildung die Folge sein könnten. Ebenso geht bei übergroßer sympathischer Erregung die Speichelproduktion zurück, was bei ängstlichen Menschen oder in Angstsituationen deutlich die Beweglichkeit der Artikulationsorgane einschränkt. Dies ist auch der Grund dafür, dass an einem Rednerpult häufig ein Glas Wasser steht. Die Emotionen, bei denen eine parasympathische Erregung auftritt, weisen dagegen eine geringere Muskelaktivität auf, weil durch sie der Herzschlag verlangsamt und die Muskulatur weniger stark durchblutet wird. Je nach Stärkegrad der parasympathischen Erregung wäre ein normaler, moderater Anspannungsgrad bis hin zu einer totalen Erschlaffung und Lähmung der laryngalen und supralaryngalen Muskulatur die Folge (Depression, Trauer).

Einstellung und Ironie

Eine bewusste Steuerung der stimmlichen und sprecherischen Eigenschaften liegt hingegen vor, wenn Sprecher/-innen in mündlichen Äußerungen durch die jeweilige Sprechweise ihre Einstellung zu Objekten bzw. Sachverhalten oder zu Personen, über die etwas mitgeteilt wird, zum Ausdruck bringen. Beschränkt man sich bei der Behandlung dieser Frage auf die prosodischen Phänomene der Sprechweise sowie auf die Dichotomie einer

positiven vs. negativen Bewertung des Mitgeteilten, so können systematische Unterschiede beobachtet werden. Die Realisierung von Sätzen wie etwa »Bayern München ist wieder Deutscher Meister geworden«, ist je nach Einstellung zu diesem Fußballverein (Fan oder Gegner) eine ganz andere; der Unterschied tritt insbesondere an der Satzmelodie der beiden Realisierungen zutage.¹⁸

Lässt man Sprecher semantisch neutrale Sätze unter Vorgabe gegensätzlicher Kontexte, die entweder eine positiv oder eine negativ wertende Sprechweise evozieren, aussprechen, so sind die korrekten Zuweisungen der Wertungen durch Hörer zu den von den Sprechern intendierten Evaluationen durchweg sehr hoch. Am höchsten fallen die korrekten Interpretationen für die positiv wertenden Sprechweisen aus: Hier liegen die Erkennungsraten zwischen 93,5% und 99,5%. Hörer sind also insgesamt sehr gut in der Lage, wertende Färbungen einer Sprechweise zuverlässig zu erkennen.

Die positiv wertende Sprechweise ist durch eine höhere mittlere Stimmlage und eine erhöhte Sprechgeschwindigkeit gekennzeichnet. Die Betonung wird ganz überwiegend primär durch Tonhöhenänderungen der Lautstärke oder Dehnungen. Die Äußerungen sind rhythmisch beschwingt und abwechslungsreich. Die negativ wertenden Äußerungen klingen hingegen häufig gelangweilt, manche erzeugen sogar den Eindruck von Unlust, Abscheu oder Ekel. Die mittlere Sprechstimmlage ist deutlich tiefer und die Sprechgeschwindigkeit deutlich langsamer. Die Stimme klingt meist normal, aber bisweilen ist sie durch Behauchungen und Laryngalisierungen gekennzeichnet. Die Betonung erfolgt primär durch Dehnungen; eine

Akzentuierung durch Tonhöhenvariation kommt dagegen selten vor. Die Satzmelodie erscheint eher monoton.

Wie aus solchen wertenden Färbungen der Sprechweise ersichtlich wird, ist die intendierte Botschaft einer sprachlichen Handlung nicht immer aus der wörtlichen Bedeutung einer Äußerung ableitbar. Eine andere Form dieses ›nicht-wörtlichen Sprechens‹ ist die verbale Ironie. Die am häufigsten vorkommende Form verbaler Ironie ist die, in der die wörtliche Bedeutung positiv und die Sprecherintention negativ ist, also eine Diskrepanz zwischen unterschiedlichen Ebenen des sprachlichen Ausdrucks vorliegt. Auf diese Form der Ironie sei hier im Hinblick auf den sprechsprachlichen Ausdruck – ebenfalls unter besonderer Berücksichtigung prosodischer Merkmale – kurz eingegangen. Wie zu erwarten, sind einige Parallelen zu den Einstellungsindikatoren zu beobachten.¹⁹

Die ironische Sprechweise unterscheidet sich von der wörtlichen durch prosodische Eigenschaften. Sie ist durch eine tiefere mittlere Grundfrequenz und eine geringere Variationsbreite charakterisiert. Die Standardabweichung der Grundfrequenz ist deutlich geringer als bei wörtlicher Sprechweise. Die Satzmelodie ist also bei ironischer Sprechweise auffallend weniger abwechslungsreich. Entsprechend zeigt die weitere Auswertung stilisierter Grundfrequenzkonturen, dass der Anteil an geraden Konturen bei der ironischen Sprechweise höher ist als bei der wörtlichen. Der Anteil an steigenden Konturen ist dagegen kleiner, und insgesamt sind die Grundfrequenzbewegungen weniger steil. Die Sprechgeschwindigkeit ist bei ironischer Sprechweise geringer. Deutliche Unterschiede sind auch in der Art der Betonung zu finden. Werden bei wörtlich ge-

meinter Sprechweise die Hervorhebungen von Silben, insbesondere des Satzakkentes, primär durch ein Tonhöhenmaximum realisiert, so wird die Hervorhebung betonter Silben in der ironischen Sprechweise vornehmlich durch eine deutliche Dehnung der Silben bei fast gleich bleibender oder sogar abgesenkter Grundfrequenz erzeugt. Interessant ist, dass die Intonationskontur der ironischen Sprechweise nicht nur eine Ähnlichkeit mit negativ wertenden Äußerungen, sondern insgesamt auch mit Konturverläufen aufweist, wie wir sie bei traurigen und langweiligen Sprechweisen im Bereich unserer Emotionsforschung gefunden haben. Diese Diskrepanz zwischen positiver Bedeutung (bei wörtlicher Interpretation) und negativ bewerteter Intonation ist wohl ein entscheidender Indikator für ironische Sprechweise.

Conclusio

In der Alltagskommunikation leiten wir unbewusst aus all den zur Verfügung stehenden Hinweisen, die uns unsere Kommunikationspartner geben, Schlüsse über diese Personen ab. Eine solche Zuordnung geschieht in der Regel ganzheitlich, spontan und intuitiv. Die Stimme und Sprechweise eines Menschen bieten hier zahlreiche Anhaltspunkte. Hörer sind in hohem Maße sensibel, derartige Informationen zu verarbeiten. Innerhalb eines Kulturkreises fällt die Beurteilung von negativen und positiven Stimmeigenschaften recht homogen aus. Allerdings sind geschlechtsspezifische Unterschiede zu berücksichtigen. Eine sympathische Frauenstimme ist nicht durch die gleichen Eigenschaften charakterisiert wie eine sympathische Männerstimme.

Für den Verlauf von Alltagskommunikation sind die Charaktereigenschaften

von großer Bedeutung. Ob ein Mensch eher besorgt und unsicher oder aber eher stabil und selbstsicher ist, kann für den Verlauf eines Gesprächs sehr entscheidend sein. Ob der Kommunikationspartner eher abenteuerlustig und gesellig oder aber eher still und nach innen gekehrt ist, kann für die Harmonie oder Dissonanz einer sozialen Interaktion ausschlaggebend sein. Aber ganz unabhängig davon, ob man einem Kommunikationspartner – beruflich oder privat – mit möglichst vielen Übereinstimmungen oder aber mit sich gegenseitig ergänzenden Eigenschaften gegenüber steht, bleibt es immer die Voraussetzung, überhaupt die Eigenschaften und die aktuelle Stimmung eines anderen Menschen einschätzen zu können, um diesen adäquat zu adressieren. Durch Stimmklang und Sprechweise findet eine Fülle zuverlässiger Informationen über die Signalquelle ihren Ausdruck. Geschickt eingesetzt sind die Möglichkeiten und Vorteile der mündlichen Rede gegenüber der Schriftsprache um ein Vielfaches höher. Wer offen und natürlich redet, dem wird Vertrauen entgegengebracht, und wer Vertrauen erhält, findet auch persönliche Anerkennung.

Abschließend sei noch auf den Ausdruck und die Wirkung von freiem Sprechen gegenüber dem Vorlesen eines Textes eingegangen. Häufig wird versucht, dem Hörenden vorzutäuschen, dass der Sprechende frei und spontan spreche – ein Versuch, der oft nicht glückt und sich in der Sprechweise deutlich bemerkbar macht. Insbesondere in der Satzmelodie sind Unterschiede zwischen einer Vorlese- und einer Sprechintonation klar herauszuhören. Die typische, ›geleitete‹ Vorleseintonation ist insbesondere durch eine zu große Zahl betonter Silben gekennzeichnet. In einem Satz sollte aber

in der Regel nur *ein* Satzakkent enthalten sein, dem die anderen Wortakzente deutlich untergeordnet sind. Nur so kann der Zuhörende die Kernaussage leicht erkennen. In freier und vielleicht sogar spontaner Rede gelingt dies Sprechenden meist mühelos und völlig unbewusst. Da sich ein Vorlesender, der in der Vortragsart des Vorlesens nicht ausgebildet ist – anders als beim freien Sprechen – an den Wörtern entlangarbeitet und nicht vorausschauend liest, wird der Sinn und damit die richtige Betonung oft erst zu spät erkannt; Betonungshäufungen sind die Folge, und die Hörenden haben so keinen unmittelbaren Zugang zum Sinn. Der häufigste Grund für das Ablesen eines vorgefertigten Textes ist die Angst vor der Sprechsituation und vor Versprechen. Zum freien Sprechen gehören die Letzteren jedoch immer dazu; sie machen sogar einen Teil der Natürlichkeit aus.

Bei dem Vorgehen, artikulatorische und akustische Korrelate für den stimmlichen Ausdruck zu bestimmen und die Relevanz einzelner Merkmale für Hörer zu erfassen, darf nie vergessen werden, dass diese zergliedernden Einzelbefunde nur dann sinnvoll interpretiert werden können, wenn sie auf die Wahrnehmung eines Sprechers / einer Sprecherin als unmittelbar gegebene und erlebte Ganzheit bezogen werden. Für Hörer gilt in der Alltagskommunikation stets das Primat der Sinn Ganzheit, zu der auch die Einbettung in einen situativen Kontext gehört. Auf dieser Basis werden von Hörern kategoriale Urteile über den stimmlichen Ausdruck von Sprechern gefällt, wobei diese nicht binär als Entweder-Oder, sondern eher als Mehr oder Weniger gefällt werden: Ein Sprecher / eine Sprecherin erscheinen aufgrund des Stimmklangs und der Sprechweise als mehr oder weniger sym-

Identitätsbilder

pathisch, mehr oder weniger glaubwürdig, mehr oder weniger aufgeregt, mehr oder weniger ärgerlich usw., wobei es sich bei solchen Attributionen stets um Wahrnehmungen handelt, wie sie ein jeweiliger Hörer unmittelbar erlebt. □

Anmerkungen

- 1 Vgl. Iris Wittlinger / Walter Sendlmeier: »Stimme und Sprechweise erfolgreicher Frauen«. In: Walter Sendlmeier (Hg.): Sprechwirkung – Sprechstile in Funk und Fernsehen. Berlin: Logos 2005, S. 71-120.
- 2 Peter R. Hofstätter / H. Lübbert: »Die Untersuchung von Stereotypen mit Hilfe des Polaritätsprofils«. In: Zeitschrift für Markt- und Meinungsforschung 3 (1958), S. 127-138.
- 3 Una M. Röhr-Sendlmeier / Barbara Schäfer / Christina Vonderbank: »Kindliche Rollenvorstellungen vom Erwachsensein im Kontext gewandelter Geschlechtsstereotype«. In: KJug (Kinder- und Jugendschutz in Wissenschaft und Praxis) 54 (2009), S. 107-111.
- 4 Walter Sendlmeier: »Mündlichkeit – Sprechstile in den Medien«. In: Walter Sendlmeier 2005, S. 1-11.
- 5 Walter Sendlmeier / Astrid Heile: »Nasalität und Behauchung als Indikatoren für den stimmlichen Ausdruck körperlichen Wohlbefagens«. In: Bernd J. Kröger / Christine Riek / Georg Sachse (Hg.): Festschrift Georg Heike, Frankfurt (M) : Hector 1998, S. 1-14.
- 6 Markus Brückl / Walter Sendlmeier: »Junge und alte Stimmen«. In: Walter Sendlmeier / Astrid Bartels (Hg.): Stimmlicher Ausdruck in der Alltagskommunikation. Berlin: Logos 2005, S. 135-163.
- 7 Sue Ellen Linville: Vocal Aging. San Diego: Singular Thomson Learning 2001.
- 8 Markus Brückl: Altersbedingte Veränderungen der Stimme und Sprechweise von Frauen. Berlin: Logos 2011.
- 9 Ralf Winkler / Walter Sendlmeier: »EEG Open Quotient in Aging Voices – Changes with Increasing Chronological Age and Its Perception«. Logopedics Phoniatrics Vocology 31 (2006), S. 51-57.
- 10 Ralf Winkler: Merkmale junger und alter Stimmen. Berlin: Logos 2008.
- 11 Gudrun Klasmeyer / Walter Sendlmeier: »Voice and emotional states«. In: Raymond D. Kent / Martin J. Ball, (Hg.): Voice Quality Measurement, San Diego: Singular Thomson Learning 2000, S. 339-357.
- 12 Walter Sendlmeier: »Phonetische Variation als Funktion unterschiedlicher Sprechstile«. In: Wolfgang Hess/Karlheinz Stöber (Hg.): Elektronische Sprachsignalverarbeitung. Bonn: Universitätsverlag 2001, S. 23-35.
- 13 Astrid Paeschke: Prosodische Analyse emotionaler Sprechweise. Berlin: Logos 2003.
- 14 Paul T. Costa / Robert R. McCrae: The NEO Personality Inventory. Manual Form S and Form R. Odessa (Florida): Psychological Assessment Resources 1985.
- 15 Ewa Stalewska: Der Ausdruck der Persönlichkeit in Stimme und Sprechweise. Mag, TU Berlin 2008.
- 16 Dagny Trägler: Der Einfluss der Synchronstimme auf die Charakterdarstellung. Mag, TU Berlin 2007.
- 17 Jan Gerhard: Die Analyse bildhafter Vorstellungen von Radiomoderatoren mit Hilfe des Programms Phantombildgenerator. Mag, TU Berlin 2008.
- 18 Ines Enterlein / Astrid Bartels / Walter Sendlmeier: »Prosodische Indikatoren der Sprechereinstellung«, In: Walter Sendlmeier / Astrid Bartels (Hg.): Stimmlicher Ausdruck in der Alltagskommunikation, Berlin: Logos 2005, S. 9-38.
- 19 Berit Johansen / Walter Sendlmeier: »Prosodische Indikatoren ironischer Sprechweise«. In: Walter Sendlmeier / Astrid Bartels (Hg.): Stimmlicher Ausdruck in der Alltagskommunikation, Berlin Logos 2005, S. 39-70.