

Decker/Kotz/Rubach

Die anwaltliche Tätigkeit in Strafsachen

Ein systematisches Praxishandbuch
für die Bearbeitung von Strafrechtsmandaten
mit Strategievorschlägen, Musterschriftsätzen und
Checklisten

20. Aktualisierungslieferung Juni 1998

Herausgegeben von Gerhard Decker,
Rechtsanwalt in Augsburg,

Dr. Peter Kotz,
Rechtsanwalt in Augsburg und

Walter Rubach,
Rechtsanwalt in Augsburg

Unter Mitarbeit von

Max Beck,
Oberstaatsanwalt a.D., Augsburg

Ulrike Liebert,
Rechtsanwältin in Augsburg

Raimund Wieser,
Richter am AG Augsburg

Dr. Andreas Wolters,
Assessor, Bremen

VERLAG
RECHT
UND
PRAXIS



Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Die **anwaltliche Tätigkeit in Strafsachen**: ein systematisches Praxishandbuch für die Beratung von Strafrechtsmandaten mit Strategievorschlägen, Musterschriftsätzen und Checklisten/ hrsg. von Gerhard Decker ... – Kissing: Verl. Recht und Praxis. –

Losebl.-Ausg.

NE: Decker, Gerhard [Hrsg.]

Grundwerk. – 1993

ISBN 3-8232-5500-2

1/4

Bearbeiterverzeichnis

Bearbeiter	Teile
Beck	8/9
Dr. Kotz	1/1 – 1/2 1/4 – 3/3 3/6 4/1 – 4/6 4/8 – 4/8.5 4/9 – 6/4.5 7/1 8/1 – 9/7
Liebert	6/4
Rubach	3/4 – 3/5 4/7.1 – 4/7.6 4/8.6 – 4/8.9 6/4.6 7/2.1 – 7/2.16
Wieser	8/5
Dr. Wolters	4/6.10

© by Verlag Recht und Praxis, Römerstraße 4, 86438 Kissing

Telefon (0 82 33) 23-4 50, Telefax (0 82 33) 23-1 47

Alle Rechte vorbehalten, Nachdruck – auch auszugsweise – nicht gestattet.

Satz: Herbert Buck, 84036 Kumhausen

Druck: Druckerei Marzorati, Königsbrunn

Printed in Germany 1998

ISBN 3-8232-5500-2

4/6.10 Beweisrechtslexikon

Das Beweisrechtslexikon dient dem raschen Auffinden von beweisrechtlich relevanten Begriffen und Problemstellungen.

Den Schlagworten ist, soweit für die Praxis bedeutsam, in dem dem Schlagwort folgenden Kapitel Rechtsprechung zugeordnet.

Zudem finden Sie, wo sinnvoll, Musteranträge etc. den Schlagworten direkt zugeordnet, um hier sofort Arbeitshilfen zur Hand zu haben.

Der Aufbau des Lexikons erfolgt alphabetisch nach Schlagworten. Innerhalb der Schlagworte liegt eine numerische Kapitelteilung vor. Das Beweisrechtslexikon wird in loser Folge ergänzt. Anregungen werden stets dankbar entgegengenommen.

Übersicht:

- AFIS (Automatisiertes Fingerabdruck-Identifizierungssystem)
- Atemalkohol
- Augenscheinsbeweis
- Betäubungsmittel, Bestimmung
- Betäubungsmittel, Bodypacking
- Betäubungsmittel, Bodystuffing
- Betäubungsmittel, Transport
- Betäubungsmittelnachweis
- Beweis Antrag
- Beweis Antrag, Ablehnungsgründe
- Beweis Antrag, Antragstellung
- Beweis Antrag, bedingter
- Beweismittel
- Beweismittel, präsent
- Beweistatsache, bedeutungslos
- Beweistatsache, erwiesene
- Beweistatsache, offenkundige
- Beweistatsache, wahrunterstellte
- Beweistatsache, zusammenhanglos
- DNA-Analyse
- Erfahrungssätze

- Erhebliche Tatsachen
- Fingerspur
- Fragerecht
- Fragetechnik
- Gegenüberstellung
 - zwecks Vernehmung (Konfrontation)
 - zwecks Wiedererkennens (Rekognition)
- Identifizieren und Wiedererkennen
- Indizienbeweis
- Lichtbildvorlage
- Sachverständigenbeweis
- Sachverständigengutachten
- Spuren
- Spurensicherung
- Spurenuntersuchung und Auswertung
- Stimmvergleich
- Textilfaserspur
- Urkundenbeweis
- Vernehmung
- Vernehmungsprotokoll
- Wiedererkennen, wiederholtes
- Zeugenaussage
- Zeugenbeweis

Stimmvergleich

1	Überblick	S. 5
2	Identifizieren von Stimmen	S. 6
2/1	Material	S. 7
2/2	Vergleichende Merkmalanalyse	S. 8
2/2.1	Personenspezifische Merkmale	S. 8
2/2.1.1	Stimmton	S. 10
2/2.1.1.1	Mittlere Sprechstimmlage	S. 11
2/2.1.1.2	Melodik der Stimme	S. 12
2/2.1.1.3	Stimmqualität	S. 13
2/2.1.2	Sprache	S. 13
2/2.1.2.1	Dialekt	S. 14
2/2.1.2.2	Akzent	S. 15
2/2.1.2.3	Idiolekt	S. 16
2/2.1.2.4	Soziolekt	S. 17
2/2.1.3	Sprechweise	S. 18
2/2.1.4	Pathologische Besonderheiten	S. 19
2/2.2	Merkmalvergleich	S. 20
2/2.2.1	Methoden	S. 20
2/2.2.1.1	Auditiv Methoden	S. 20
2/2.2.1.2	Apparative Methoden	S. 21
2/2.2.1.3	Kombinierte Methoden	S. 23
2/2.2.1.4	Methoden technischer Aufbereitung	S. 24
2/2.2.2	Dauer der Untersuchung	S. 25
2/3	Befund	S. 25
2/4	Befundbewertung	S. 26
2/5	Fehlerquellen	S. 30
2/5.1	Material	S. 30
2/5.1.1	Zeitlicher Umfang der Tonaufzeichnung	S. 30
2/5.1.2	Telephonaufzeichnung	S. 30
2/5.1.3	Akustische Besonderheiten	S. 31
2/5.1.3.1	Verstellung	S. 31
2/5.1.3.2	Natürliche Stimmprofilveränderung	S. 32
2/5.1.3.3	Sprechsituation	S. 33
2/5.2	Methode	S. 34
2/5.2.1	Technische Bearbeitung des Spurenmaterials	S. 34
2/5.2.2	Korrekte Meßvoraussetzungen	S. 35
2/5.2.3	„Subjektiver Faktor“	S. 35
2/5.2.4	Visualisierung	S. 36
2/5.3	Erfahrungswissen	S. 37

2/5.4	Bewertung	S. 37
2/6	Resümee	S. 37
3	Wiedererkennen von Stimmen	S. 39
3/1	Hilfsmittel	S. 39
3/2	Typisierendes Erleben	S. 40
3/3	Subjektive Behauptung	S. 42
3/4	Ergebniskontrolle	S. 43
3/5	Fehlerquellen	S. 45
3/5.1	Wahrnehmung	S. 45
3/5.2	Erinnerung	S. 47
3/5.3	Wiedergabe	S. 48
3/5.4	Gegenüberstellung	S. 49
3/5.4.1	Voraussetzungen	S. 50
3/5.4.2	Vergleichsstimmen	S. 51
3/5.4.3	Vergleichstext	S. 52
3/5.4.4	Personal	S. 53
3/5.4.5	Ablauf	S. 53
3/5.4.6	Wiederholung	S. 54
3/5.4.7	Dokumentation	S. 54
3/6	Beweiswert	S. 55
3/7	Rechtsprechung	S. 56
3/7.1	Wahlgegenüberstellung	S. 56
3/7.2	Dokumentation	S. 57
3/7.3	Beweiswert	S. 57
3/7.4	Beweisantrag	S. 58
3/8	Resümee	S. 59
4	Übersicht	S. 60

Literatur¹:

Allgemein

Brinker, Identifizieren und Wiedererkennen – Bemerkungen zum Unterschied und zur Beweisqualität, ArchfKrim 176 (1985), 142

Identifizieren von Stimmen

Braun/Broeders, Zu den Möglichkeiten einer technischen Qualitätsverbesserung von Tonaufzeichnungen, NStZ 1996, 173

¹ Kurzbelege im Text, z.B. [ArchfKrim 176,142], verweisen auf die angegebenen Seiten nachfolgender Literatur einschließlich Fremdzitaten und Fußnoten.

Bunge, Möglichkeiten und Grenzen der Sprechererkennung und der Tonbandauswertung für die polizeiliche Praxis, ArchfKrim 176 (1985), 65

Burghard u.a., Kriminalistik Lexikon, 3. Aufl. 1996

Krause, Verfahren zum Aufdecken von Tonbandmanipulationen, ArchfKrim 172 (1983), 153

Krause, Möglichkeiten und Grenzen der Sprecheridentifizierung, ArchfKrim 157 (1976), 154

Kriminalistik International, Sprechererkennung – beeinflusst von Sprechgeschwindigkeit und Lautstärke?, Kriminalistik 1996, 118
Künzel, Forensische Phonetik – Beispiel einer Symbiose von Geisteswissenschaft und Hochtechnologie, ArchfKrim 194 (1994), 129

Künzel, Die Erkennung von Personen anhand ihrer Stimme, in: Kube/Störzer/Timm (Hrsg.), Kriminalistik, Bd. 1 (1992), 817

Künzel, Die Erkennung von Personen anhand ihrer Stimme, NStZ 1989, 400

Künzel, Dem Täter auf der Stimmspur – Praxis der forensischen Sprechererkennung, Kriminalistik 1985, 120

Masthoff, Automatische Sprecherifizierung bei Telefonübertragung, FS Wängler (1987), 223

Moeller, Identifizierungshilfe „Stimme“ – Ein „neuer“ Ermittlungsansatz zur Diskussion gestellt, Kriminalistik 1993, 588

van der Giet/Künzel, Rechnergestützter Stimmvergleich für forensische Anwendungen, Kriminalistik 1981, 341

Wiedererkennen von Stimmen

Artkämper, Sofortfahndung und Wiedererkennen – Einzelgegenüberstellung und richterliche Überzeugungsbildung, Kriminalistik 1997, 505

Artkämper, Gegenüberstellungen – Erkenntnisquelle mit Kautelen (Zu einigen praktischen Problemen im Umgang mit Identifizierungsmaßnahmen im Sinne des § 58 Abs. 2 StPO), Kriminalistik 1995, 645

Bender/Nack, Tatsachenfeststellung vor Gericht, Bd. 2 (Vernehmungsslehre), 2. Aufl. 1995

Damm, Polizistenmord in Holzminden, Kriminalistik 1996, 119, 189

Eisenberg, Beweisrecht der StPO, 2. Aufl. 1996

Eisenberg, Visuelle und auditive Gegenüberstellung im Strafverfahren – empirische und rechtliche Zusammenhänge, Kriminalistik 1995, 458

Glaser, Wiedererkennen im Rahmen der Sofortfahndung – Auswege aus einem anscheinend unauflösbaren Dilemma, Kriminalistik 1995, 653

Görling, Täteridentifizierung per Videogegenüberstellung, Kriminalistik 1985, 58

Hammersley/Read, Das Wiedererkennen von Stimmen, in: Köhnken/Sporer (Hrsg.), Identifizierung von Tatverdächtigen durch Augenzeugen, 1990

Köhnken, Gegenüberstellungen – Fehlerquellen bei der Identifizierung durch Augenzeugen, Kriminalistik 1993, 231

Künzel, Stimmen-Identifizierung im Fall Dr. Schneider – Probleme bei der Sprecher-Identifizierung durch Laien, ArchfKrim 198 (1996), 1

Künzel, Zum Problem der Sprecheridentifizierung durch Opfer und Zeugen, GA 1988, 215

Nöldeke, Zum Wiedererkennen des Tatverdächtigen bei Gegenüberstellung und Bildvorlage, NStZ 1982, 193

Steinke, Der zweifache Polizistenmord von Holzminden, Kriminalistik 1996, 287

Undeutsch, Die Wiedererkennung von Personen, FS Peters (1984), 461

1 Überblick

In den folgenden Ausführungen geht es um die kriminalistische Bedeutung der menschlichen Stimme im Rahmen des strafprozessualen Identitätsnachweises (siehe dazu den Fragenkatalog „Identifizieren und Wiedererkennen“).

Zu prozeßrechtlichen Fragestellungen im Zusammenhang mit dem kriminalistischen Stimmvergleich siehe auch Teil 3/6 (Überwachung des Fernmeldeverkehrs), 3/3.10 (Hörfallen) und 3/2 (Akteneinsicht).

Spezielle Probleme und Techniken der „Tonbandauswertung“ bleiben einem gesonderten Stichwort vorbehalten.

2 Identifizieren von Stimmen

„Identifizieren“ wird hier im Sinne eines Identitätsnachweises aufgrund „objektbezogener Analyse“ verstanden (siehe „Identifizieren und Wiedererkennen“).

Objektbezogene Analysen werden als Vorgänge des Verstehens beschrieben, die „objektive Gewißheit“ in einem Überzeugungsprozeß intendieren, der das „Nachvollziehenkönnen“ durch andere ermöglichen soll [ArchfKrim 176, 143] („intersubjektive Überprüfbarkeit“).

Wer menschliche Stimmen objektbezogen analysiert, ist mit seinem analytischen Verstand (Intellekt) gefordert, der mit gezielter Aufmerksamkeit das Vordergründige ganzheitlicher Erscheinungen durchdringt, um sich dem Detail, einem Merkmal oder einem Merkmalkomplex zuzuwenden [ArchfKrim 176,143].

Das Identifizieren von Stimmen ist ein Arbeitsfeld der forensischen Phonetik, einer auf forensische Bedürfnisse zugeschnittenen sprachwissenschaftlichen Fachdisziplin, die sich in der Bundesrepublik Deutschland in den frühen achtziger Jahren herausgebildet hat [ArchfKrim 194,129]. In dem „Streben nach Objektivierung kriminaltechnischer Untersuchungsergebnisse zum Zwecke der Verwendung im Gerichtsverfahren mit nicht-kooperativen Angeklagten“ wird dabei in der Bundesrepublik Deutschland „von Anfang an großer Wert auf eine rechnergestützte, apparative Komponente gelegt“ [ArchfKrim 194,130]. „Die visionäre Zielsetzung von Horst Herold, dem Präsidenten des Bundeskriminalamtes von 1971 bis 1981, war die Ersetzung des Zeugen durch den Computer: Wo immer möglich, sollte der als subjektiv und irrtumsbehaftet angesehene Zeugenbeweis vom Sachbeweis abgelöst werden. Die als Konsequenz dieses Gedankens im Bundeskriminalamt stark geförderte rechnergestützte Phonetik ist daher in den achtziger Jahren eine technologische Speerspitze der Phonetik schlechthin gewesen“ [ArchfKrim 194,130] (zur „täuschend suggestiven“ Wirkung der herkömmlichen Unterscheidung zwischen „Personalbeweis“ und „Sachbeweis“ siehe „Sachverständigengutachten“, „Indizienbeweis“, „Spurenuntersuchung und Auswertung“ sowie „Identifizieren und Wiedererkennen“).

Die Zielsetzung einer vollautomatischen Durchführung forensischer Stimmvergleichsanalysen („wie eine Atemalkoholanalyse“ [!?!]) gilt mittlerweile als utopisch, als „realistisch“ betrachtete Projekte wie die Entwicklung eines interaktiven Mensch-Rechner-Systems wurden zwischenzeitlich aufgegeben oder „storniert“ [NStZ 1989, 405].

2/1 Material

Als „forensische Sprachaufzeichnungen“ [NStZ 1989, 401] kommen Spurenmaterial und zu Vergleichszwecken erstellte Sprachproben in Betracht.

Im Gegensatz zu Sprachaufzeichnungen kommerzieller Stimmvergleichssysteme etwa im Bereich von Zugangskontrollen geschützter Objekte, bei denen „mit Hilfe einer festgelegten sprachlichen Äußerung eine gewünschte ausweisähnliche Identifizierung erzielt werden soll“ („Sprecherverification“) [ArchfKrim 157, 154], ist forensisch bedeutsames Sprachmaterial in aller Regel durch folgende Besonderheiten gekennzeichnet [NStZ 1989, 401]:

- keine Kooperation der sprechenden Person (die Person möchte nicht erkannt werden, verstellt ihre Stimme oder ahmt fremde Stimmen nach [Kriminalistik 1981, 341]);
- keine textidentische Vergleichsbasis (Spurenmaterial und Vergleichsproben haben in aller Regel unterschiedliche Sprachinhalte);
- akustische Störquellen (resultierend aus Sprechsituation und Aufnahmetechnik);
- Spuren- und Vergleichsmaterial können identisch sein, müssen es aber nicht (die potentielle Vergleichsmenge ist groß und offen).

Für die Eignung „akustischen Spurenmaterials“ („ASM“) zum Zwecke kriminaltechnischer Untersuchung sind insbesondere dessen technische Qualität und zeitliche Dauer von Bedeutung [Burghard 8]. Tonaufzeichnungen von nur 5 Sekunden Dauer sind in der forensischen Praxis keine Seltenheit; eine Dauer von 60 Sekunden ist bereits als überdurchschnittlich anzusehen [NStZ 1989, 402].

Für die optimale forensische Verwertbarkeit von Sprachaufzeichnungen sind folgende Voraussetzungen zu beachten [Burghard 296]:

- Aufzeichnung nur auf gute (relativ teure) Spulen- oder Kassettengeräte;
- Gerätebetrieb nur über Netzstrom, nicht über Batterie;
- Verwendung fabrikneuer Langspielbänder oder C-60-Kassetten;

- korrekter Anschluß an Fernmeldegeräte (Postanschlußglied);
- korrekte Aussteuerung, erforderliche Korrekturen per Hand;
- korrekte Bandgeschwindigkeit (nach Möglichkeit 9,5 cm/sec, 4,75 cm/sec nur in Notfällen, 2,4 cm/sec in keinem Fall) [ArchfKrim 176, 79];
- Vergleichsproben in ruhiger Räumlichkeit sowohl in freier Rede als auch mit vorbereitetem Text;
- Verlesung von Texten unter 30 Sekunden Dauer zehnmal, unter einer Minute Dauer fünfmal und längere Texte zweimal;
- für Spuren- und Vergleichsproben möglichst gleiche Tonbandgeräte benutzen;
- ist der zur Tatausführung benutzte Telefonanschluß bekannt, nach Möglichkeit von dort aus sprechen lassen.

„Akustisches Vergleichsmaterial“ („AVM“) [Burghard 8] kann auch „u.U. aus den spontanen Äußerungen eines Sprechers, beispielsweise eines Angeklagten oder Zeugen in der Hauptverhandlung, gewonnen werden“ [GA 1988, 216]¹.

¹ Zum Problem der „Hörfallen“ und anderer Methoden legaler und illegaler Beschaffung von Vergleichsmaterial siehe Teil 3/3.10.

2/2 Vergleichende Merkmalanalyse

Der kriminaltechnische Nachweis einer Identität von Stimmen erfolgt durch vergleichende Merkmalanalyse.

2/2.1 Personenspezifische Merkmale

Grundlage vergleichender Stimmanalysen ist das lautsprachliche (verbale) Verhalten einer Person [NStZ 1989,400], das im Hinblick auf individualisierende Merkmale hin untersucht und ausgewertet wird.

Dabei werden „sprecherspezifische Merkmale“ und „sprachsystemspezifische“ Merkmale (Bedeutungsgehalt von Sprachlauten und größeren Einheiten) unterschieden (und zwar vor allem im Hinblick auf eine Automatisierung der Stimmvergleichsanalyse [NStZ 1989, 401]).

Zur Kennzeichnung sprachlicher Phänomene als „sprecherspezifische“ Merkmale werden unter den Bedingungen forensisch bedeutsamen Sprachmaterials vor allem drei Teilbereiche sprachlicher Kommunikation ausgewertet [NStZ 1989,400]: der Stimmton (als im menschlichen Kehlkopf erzeugte Schwingung), die Sprache (als akustische Ausprägung der jeweiligen Landessprache) und die Sprechweise (als gewohnheitsmäßige Stimmbildung und Lautproduktion).

Die für die forensische Stimmvergleichsanalyse bedeutsamen Merkmale und Merkmalskomplexe sind in der folgenden Übersicht zusammengestellt; die Begriffe „sprecherspezifisch“ und „personenspezifisch“ werden hier gleichbedeutend verwendet.

Übersicht: Personenspezifische Merkmale [NStZ 1989, 403]

Lautsprachliches		
Verhalten	Merkmal oder Merkmalskomplex	Darstellung
Stimmton	Mittlere Sprechstimmlage Melodik der Stimme Stimmqualität	Mittelwert der Grundfrequenz Streuungsmaß der Grundfrequenz physikalische Form der Schwingungen
Sprache	Dialekt	Art und Grad der dialektalen Färbung fremdsprachliche Interferenzen
	„fremdsprachiger Akzent“ Idiolekt einschließlich Sterotypen im Bereich des Wortes oder größerer Einheiten	Aspiration von Verschlusslauten Nasalität akustische Realisierung Verteilungseigenschaften von Verzögerungen (Häsitationen)
	Soziolekt	
Sprechweise	Artikulationsgeschwindigkeit Atemverhalten	Sprechtempo

Für die Eignung „akustischen Spurenmaterials“ zum Zwecke eines kriminaltechnischen Identitätsnachweises sind neben der technischen Qualität und zeitlichen Dauer der Tonaufzeichnung insbesondere „die Anzahl, Struktur und Stabilität der vergleichbaren sprecherindividualisierenden Merkmale“ von Bedeutung [Burghard 8]. Die personenspezifische Potenz sämtlicher Merkmale nimmt in dem Maße zu, wie sie von durchschnittlichen Werten abweicht [NStZ 1989, 404]. Bei allen Merkmalen (außer dem Dialekt) können zudem pathologische Ausprägungen auftreten (Stimm-, Sprach- und Sprechstörungen), die eine erhebliche Steigerung ihres personenspezifischen Wertes bewirken [Künzel 826].

Im folgenden werden einzelne Merkmalskomplexe skizzenhaft erläutert.

2/2.1.1 Stimmton

Ein Stimmton wird dadurch erzeugt, daß zwei im Kehlkopf zwischen verschiedenen Knorpeln nebeneinander aufgespannte Stimmbänder durch den aus der Luftröhre entweichenden Luftstrom in Schwingung versetzt werden [Künzel 823]. Physikalisch betrachtet hat die erzeugte Schwingung „eine etwa sägezahnartige Form“ und ist außer ihrer Grundfrequenz durch eine große Anzahl von Oberschwingungen charakterisiert [ArchfKrim 176, 71].

Bei der Erzeugung stimmloser Laute sind die Stimmbänder unbeteiligt: Durch Zungenspitze, Lippen und Gaumen werden enge Spalten gebildet, an denen die mit hoher Geschwindigkeit vorbeiströmende Luft durch Turbulenzbildung Rauschtöne erzeugt [ArchfKrim 176, 71].

Bei der lautsprachlichen Artikulation wird „dieses Schwingungsgemisch“ durch Öffnen und Schließen des Mundes, durch die Lage der Zunge sowie der An- und Abkoppelung des Nasenraumes durch das „Velum“ (Gaumensegel) laufend geändert; auf diese Weise können jeweils unterschiedliche Oberschwingungen resonanzmäßig in der Amplitude (größter Ausschlag einer Schwingung) angehoben („Formanten“) oder abgesenkt („Antiformanten“) werden [ArchfKrim 176, 71].

2/2.1.1.1 Mittlere Sprechstimmlage

Die Höhe des Stimmtons hängt von der Schwingungsfrequenz der Stimmbänder ab; je schneller die Stimmbänder schwingen, desto höher wird der erzeugte Stimmton von einem Zuhörer empfunden [Künzel 823].

Als „mittlere Sprechstimmlage“ wird in der Medizin („Phoniatry“) die „Tonhöhe“ bezeichnet, „von der für jeweils kurze Zeit die Satzmelodie nach oben und unten abweicht“ (Grundfrequenz der Stimmbänder) [NStZ 1989, 403].

Physikalisch wird die Höhe des Stimmtons in Hertz (Hz) gemessen; diese Maßeinheit bezeichnet die Anzahl der Schwingungen pro Sekunde (Stimmtonfrequenz) [Künzel 823].

Empirische Untersuchungen an großen Personengruppen vermitteln Erkenntnisse über eine Gesamtverteilung der Stimmbandgrundfrequenzen männlicher und weiblicher Erwachsener („durchschnittliche Stimmfrequenz“) und finden über Hintergrundstatistiken Eingang in die forensische Begutachtung [NStZ 1989, 403] (Erfahrungsbasis!).

Beispiel: Die Stimmbandgrundfrequenz liegt im Mittel bei Männern zwischen 70 und 150 Hz, bei Frauen zwischen 200 und 300 Hz; die durchschnittliche Höhe der Stimme deutscher männlicher Erwachsener beträgt 116 Hz (Mittelwert) [ArchfKrim 198, 3] (Frauen und Kinder haben im allgemeinen höhere Stimmen als Männer, da ihr Kehlkopf kleiner und die Stimmbänder entsprechend kürzer sind und infolgedessen auch schneller schwingen [Künzel 823]).

Da das Schwingverhalten der Stimmbänder von einer Vielzahl anatomischer, physiologischer und konstitutioneller Faktoren abhängt, gilt die Grundfrequenz als ein besonders typisches personenbezogenes Merkmal, das darüber hinaus als äußerst robust gegenüber Verzerrungen des Telefonkanals und verhältnismäßig zuverlässig ableitbar ist; allerdings kann sie von der Sprechsituation, dem psychologischen Status der Sprechenden Person (wie etwa Streß), der Anwesenheit weiterer Gesprächspartner und anderen Faktoren abhängen [Kriminalistik 1981, 345].

Nach einer neueren Feldstudie besteht zwischen der Hauptgrundfrequenz („informativer akustischer Parameter“) und anderen Merkmalen wie Sprechgeschwindigkeit und Lautstärke keine statistisch signifikante Beziehung; vor Abschätzung und Beurteilung der Grundfrequenzmittelwerte ist deshalb eine Anwendung von „Korrektur- und Normalisierungsverfahren“ nicht erforderlich [Kriminalistik 1996, 118].

Das Merkmal „mittlere Sprechstimmlage“ wird als eines der am besten erforschten bezeichnet [NStZ 1989,403]; seine Bestimmung erfolgt „weitgehend“ automatisiert [NStZ 1989, 404].

2/2.1.1.2 Melodik der Stimme

Der Stimmton kann beim Sprechen durch komplizierte muskuläre Mechanismen auf vielfältige Weise verändert werden, wobei

der Verlauf des Stimmtons im Satz (Intonation, Sprechmelodie) auch zur Übermittlung bestimmter Aspekte der sprachlichen Nachricht benutzt wird [Künzel 823].

Beispiel: Ein der Wortstellung nach als Aussage erscheinender Satz kann allein durch Anheben des Stimmtons gegen Ende als Frage gekennzeichnet werden: Der Täter wurde wirklich festgenommen (?); Befehle werden im allgemeinen durch starkes Absenken des Stimmtons markiert [Künzel 823].

Technisch wird die Sprechmelodie als von der mittleren Grundfrequenz abgeleitete Streuung erfaßt (mathematische Standardabweichung): Eine geringe Standardabweichung bezeichnet eine monotone Stimme, eine große Standardabweichung eine stark modulierte Stimme [NStZ 1989, 403].

Individualtypisch kann die Verteilungsform der Stimmtonabweichung sein, wenn etwa statt kontinuierlicher Veränderungen bestimmte Sprünge (Intervalle) benutzt werden [NStZ 1989,403].

Der Verlauf des Stimmtons kann im übrigen auch Hinweise auf Geschlecht, Alter, Trunkenheit, Angst, Ironie und andere „paralinguistische Merkmale“ vermitteln [Künzel 823].

2/2.1.1.3 Stimmqualität

Der für eine menschliche Stimme charakteristische Klang ergibt sich aus der geometrischen Form, Beschaffenheit und Größe der Hohlräume des Mundes, der Nase und des Rachens, die oberhalb des Kehlkopfes als Resonanzräume wirken [Künzel 824].

Aufgrund von Unterschieden in der Gestalt der Stimmbandschwingungen kann sich der Klang einer Stimme in vielfältiger Weise verändern [NStZ 1989, 403].

Beispiel: Bei nicht völlig luftdichtem Verschuß an den Stimmbändern entsteht auditiv ein Eindruck von Heiserkeit; bei extrem tiefer Frequenz und daher nicht synchronem Schwingen der Stimmbänder entsteht die sogenannte Knarrstimme [NStZ 1989,403].

Im Einzelfall vorliegende Besonderheiten werden zunächst nach medizinischen Kriterien diagnostiziert und mit Hilfe apparativer Verfahren dargestellt (Kurzzeit- und Langzeitspektrographie) [NStZ 1989, 403].

Für die forensische Anwendung gilt dieser Merkmalbereich allerdings als nahezu wertlos, weil er in ausgeprägter und überwiegend nicht kompensierbarer Weise von den Einflüssen der Telefonübertragung beeinträchtigt wird [Künzel 824].

2/2.1.2 Sprache

Sprache wird dadurch erzeugt, daß die „Hohlraumkonfiguration“ oberhalb der Stimmbänder (Rachen, Gaumen, Nase, Mund) mit oder ohne Stimmbandschwingung verändert wird; durch zielgerichtete artikulatorische Bewegungen (artikulatorische Gesten) wird ein akustisches Signal erzeugt, das von einem Hörer in bedeutungstragende lautliche Einheiten zerlegt wird [Künzel 824].

Beispiel: Durch kurzzeitiges Verschließen des Mundraumes oberhalb der oberen Schneidezähne wird mit Hilfe der Zungenspitze der Laut „t“ erzeugt und von einem Hörer wahrgenommen; bei zusätzlichem Schwingen der Stimmbänder wird der Laut „d“ erzeugt und wahrgenommen [Künzel 824].

Das Zusammenwirken der Sprechwerkzeuge (Artikulationsorgane) im einzelnen ist ein äußerst komplexes Geschehen, wobei insbesondere auch zu berücksichtigen ist, daß fortlaufendes Sprechen keine bloße Aneinanderreihung abgegrenzter Sprachlaute bzw. artikulatorischer Bewegungen darstellt, sondern daß Sprachlaute bzw. Gesten sich in hohem Maße gegenseitig beeinflussen („Koartikulation“) [Künzel 824].

Die wissenschaftliche Forschung systematisiert die Sprachlaute nach verschiedenen Klassen. Grundlegend ist dabei die Unterscheidung zwischen Konsonanten und Vokalen: Ist bei der Artikulation des Sprachlautes ein Stimmton vorhanden, ist der Sprachlaut stimmhaft (Vokal), ohne Stimmton ist er stimmlos (Konsonant) [Künzel 825].

Konsonanten werden nach artikulatorischen (auf ihrer Erzeugung beruhenden) Kriterien klassifiziert [Künzel 825].

Beispiel: bilabial (Unterlippe, Oberlippe) für „p“ und „b“; labiodental (Unterlippe, obere Schneidezähne) für „f“ und „v“; dorsovelar (Zungenrücken, weicher Gaumen) für „k“ und „g“ [Künzel 825].

Vokale werden mangels hinreichend präziser Beschreibungsmöglichkeiten der jeweiligen Artikulationsbewegungen mittels perzeptorischer (an der Wahrnehmung orientierter) Kategorien klassifiziert; als Referenzsystem für sämtliche Vokalartikulationen dient ein System sogenannter „Kardinalvokale“, welches der Fachwissenschaftler (Phonetiker) sich in gründlichem Training aneignen muß [Künzel 825].

Vokale und Konsonanten können auch mit Hilfe apparativer Verfahren, insbesondere der Spektralanalyse, beschrieben werden [Künzel 825].

Als personenspezifisch „potente“ Sprachlaute gelten insbesondere „a“, „m“ und „n“ [NSTZ 1989,402].

2/2.1.2.1 Dialekt

In der Bundesrepublik Deutschland werden mehr als fünfzig verschiedene Dialekte und Mundarten gesprochen, die sich generell in den verschiedenen Aussprachekategorien einzelner Phoneme (kleinste bedeutungsunterscheidende, aber nicht selbst bedeutungstragende sprachliche Einheiten), der Intonation von Phonemsequenzen als auch durch die Verwendung sprachraumbezogener Worte und Floskeln unterscheiden [ArchfKrim 176, 69].

Als regionales Eingrenzungskriterium gilt der Dialekt als ein wichtiges personenspezifisches Merkmal des lautsprachlichen Verhaltens; dabei werden nicht nur die Art des Dialekts sondern auch der Grad der dialektalen Färbung unterschieden [NSTZ 1989, 403].

Charakteristische Besonderheiten werden von dialektologisch geschulten Sprachwissenschaftlern mit Hilfe ihres „wissenschaftlichen Instrumentariums“ und ihrer durch jahrelanges systematisches „Perzeptionstraining“ geschulten Ohren erfaßt und können mit Hilfe apparativer Verfahren (insbesondere der Spektrographie) „objektiviert“ werden [NSTZ 1989, 403].

Die Grenzen einer wissenschaftlichen Dialektbestimmung sind dann erreicht, wenn die fragliche Person in dem sprachprägenden Lebensalter längere Zeit in unterschiedlichen Sprachräumen gelebt hat und daher einen „Mischdialekt“ mehr oder weniger starker Ausprägung spricht, oder wenn ein überregionales unspezifisches Hochdeutsch gesprochen wird [ArchfKrim 176, 69].

Im Rahmen eines im Auftrag des Bundeskriminalamtes von den Universitäten Marburg und Trier durchgeführten Projektes „Datenbank regionaler Umgangssprachen (DRUGS)“ wird empirisch erhobenes Sprachmaterial aus allen Teilen des deutschen Sprachgebietes nach dialektologischen, phonetischen, phonologischen und morphologischen Kriterien analysiert und anschließend nach forensisch-praktischen Gesichtspunkten so strukturiert, daß Benutzer, die nicht unbedingt Germanisten oder Dialektologen sein müssen, mittels gezielter CD-ROM-Recherche eigene Analyseergebnisse im konkreten Fall überprüfen können [ArchfKrim 194, 140].

2/2.1.2.2 Akzent

Liegen fremdsprachliche Äußerungen vor, kann die Nationalität der Sprechenden Person in aller Regel problemlos ermittelt werden [ArchfKrim 176, 68].

Demgegenüber ist die nationale Herkunft beim sogenannten „Ausländerdeutsch“ schwieriger zu bestimmen, da hier eine Mischung zweier Sprachen mit unterschiedlichen Anteilen vorliegt [ArchfKrim 176,68]. Zur Ermittlung eines möglicherweise fremdsprachlichen Akzents werden im Rahmen der „auditiven Analyse“ durch einen „kontrastiven phonologisch-phonetischen Vergleich zweier Lautsysteme“ etwaige muttersprachlich bedingte Ausspracheabweichungen („Interferenzen“) ermittelt [ArchfKrim 194, 133]. Voraussetzung für eine exakte Nationalitätsbestimmung ist die Kenntnis des entsprechenden fremdsprachlichen Laut- und Intonationssystems [ArchfKrim 176, 68].

Die „phonetische Differentialdiagnostik“ kommt allerdings in der Praxis nicht selten zu dem Ergebnis, daß die Sprechende Person gar keine fremde sondern Deutsch als Muttersprache besitzt und das, was sie beispielsweise für einen italienischen Akzent hält, nur vortäuscht; für einen Fachmann ist solch eine Akzent-

vortäuschung („diffuses Gastarbeiterdeutsch“) leicht an der „Inkonsistenz des fremdsprachlichen Anteils“ [ArchfKrim 176,68] erkennbar, da sich die sprechende Person in der Regel nicht bewußt ist, daß Portugiesen, Italiener oder Türken aufgrund ihres muttersprachlichen Lautsystems völlig verschiedenartige „Fehler“ begehen, wenn sie Deutsch sprechen [ArchfKrim 194, 133].

Beispiel: Im Gegensatz zu Engländern oder Iranern wird eine Person mit türkischer Muttersprache mit der Aussprache von 'ü' und 'ö' keinerlei Probleme haben [ArchfKrim 194, 133].

2/2.1.2.3 Idiolekt

Als Idiolekt werden sprachliche Merkmale bezeichnet, die nicht dialektaler Art sind sondern einen „individuellen Habitus“ darstellen [NStZ 1989, 403].

Beispiel: Wiederholter Gebrauch bestimmter Redewendungen und Floskeln (Stereotypen) wie „nicht wahr“, „sozusagen“, „eben halt“ oder exotischer Floskeln wie „nicht wahr nicht“, „jawoll“, „jawoll“ (mit besonderer Intonationstruktur) [NStZ 1989, 403].

Weitere idiolektische Merkmale sind etwa Art und Häufigkeit von Versprechern, bestimmte Formen der Nasalität, oder die Häufigkeit, Verteilung und phonetische Realisierung von Pausen und Verzögerungen [NStZ 1989, 403]; dazu gehören auch „Hesitationserscheinungen“, die von kaum wahrnehmbaren Sprechhemmungen und Verzögerungen bis hin zum Stottern reichen, eine anomale Zeit- und Pausenstruktur oder eine besondere Häufung von redundanten Füllsilben wie „äh“, „ähm“ oder „hm“ [ArchfKrim 176, 69].

Als äußerst potentes Merkmal gilt die „Artikulationsrate“ als Anzahl der pro Zeiteinheit durchschnittlich produzierten Silben abzüglich aller Pausen und Verzögerungen; dieses Merkmal erweist sich gegenüber Unterschieden in der Sprechsituation als äußerst robust [Kriminalistik 1985, 122].

Idiolektische Merkmale werden vom Sachverständigen zunächst subjektiv erfaßt und können mit Hilfe apparativer Verfahren dargestellt werden [NStZ 1989, 403].

2/2.1.2.4 Soziolekt

Als Soziolekt werden schichtenspezifische Merkmale bezeichnet [NStZ 1989, 403].

Beispiel: Unterschiede in der sprachlichen Artikulation von Arbeitern und Akademikern oder Berufsgruppen, bei denen die sprachliche Kommunikation eine wesentliche Rolle spielt, und die sich teilweise besonderer Fachsprachen (Jargon) bedienen [NStZ 1989,403].

In der forensischen Praxis dienen soziolektale Merkmale häufig der Eingrenzung des Täterkreises [NStZ 1989, 403].

Beispiel: Aufgrund grammatikalischer Merkmale wie Komplexität des Satzbaus oder Wortwahl kann in aller Regel die Zugehörigkeit zu einem „sprechenden“ Beruf (Lehrer, Rechtsanwälte, Journalisten oder Büroangestellte) oder einem „nichtsprechenden“ Beruf (Handwerker, Landwirt, Hilfsarbeiter) unterschieden werden [NStZ 1989, 403].

Präzise Angaben aufgrund soziolektischer Merkmale erfordern jedoch eine ausreichende Menge an auswertbarem Sprachmaterial [NStZ 1989, 403].

2/2.1.3 Sprechweise

Die Sprechweise eines Menschen resultiert aus der Art und Weise, wie er gewohnheitsmäßig Stimmbildung und Lautproduktion in der sprachlichen Kommunikation einsetzt [Künzel 825]. Von ausschlaggebender Bedeutung sind seine generellen und augenblicklichen stimmlichen, artikulatorischen und rhetorischen Fähigkeiten, aber auch seine Einstellung zum Kommunikationsereignis, seine Motivation und Emotion sowie sein soziales Verhalten und der Einfluß des sozialen und akustischen Umfeldes [Burghard 299].

Als personenspezifische Merkmale gelten insbesondere die Sprechgeschwindigkeit, Betonung (Muster von betonten und unbetonten Sprechphasen), Pausen und Verzögerungsmerkmale (Häufigkeit, Verteilungsmuster und Art) sowie das Atemverhalten; sämtliche Merkmale können „mit Hilfe subjektiver und apparativer Verfahren“ klassifiziert werden [Künzel 825].

Zur Sprechgeschwindigkeit wird auf empirische Untersuchungen des Bundeskriminalamtes hingewiesen, wonach die Anzahl der pro Zeiteinheit produzierten Silben abzüglich aller Pausen und Verzögerungen ein personenspezifisch relativ stabiles Maß ergibt („articulation rate“), während sich Sprecher untereinander deutlich unterscheiden; für die Erstellung einer Hintergrundstatistik reichen die vorhandenen Datenmengen (Versuchspersonen und forensische Fälle) noch nicht aus [Künzel 829].

Das Atemverhalten kann hinsichtlich Rhythmus und Dauer der Atemzüge sowie ihres Geräusches ausgewertet werden; auch insoweit ist zunächst „das subjektiv diagnostische Vermögen“ des Sachverständigen gefragt, eine „Objektivierung“ erfolgt durch verschiedene Meßverfahren [Künzel 829].

Beispiel: Spektralanalyse bezüglich des Atemgeräusches aufgrund einer Verengung des Querschnitts der Luftröhre oder des Kehlkopfes infolge Bronchitis oder akute Heuschnupfens [NStZ 1989, 404].

Zum Merkmalkomplex Sprechweise gehören auch gewisse nichtsprachliche Angewohnheiten.

Beispiel: Auffälligkeiten im Atemverhalten, habituelles Schniefen oder Räuspfern sowie bestimmte Geräusche beim Öffnen der Lippen („Clicks“ oder Schmatzlaute) [NStZ 1989, 404].

Derartige Umstände können auch Hinweise auf die emotionale Befindlichkeit der Sprechenden Person geben [Künzel 818].

2/2.1.4 Pathologische Besonderheiten

Lautsprachliche Besonderheiten können sich auch aus körperlichen Ursachen ergeben, wie etwa einem Spaltrachen („Hasenscharte“), Polypen, Stimmbandveränderungen (Knötchen, Krebs), aber auch durch Zahnstandsregulierungen und Zahnimplantate oder Zahnprothesen [ArchfKrim 176, 69].

Beispiele: Das Tragen einer Vollprothese am Oberkiefer oder einer Brücke an den oberen Schneidezähnen verändert möglicherweise den Klang bestimmter Reibelauten oder es tritt als

Ergebnis der veränderten taktilen und propriozeptiven (das Bewegungsgefühl betreffenden) Rückmeldung der dentalen oder alveolaren (an oder hinter den Zähnen befindlichen) Artikulationspunkten ein Lispeln auf [ArchfKrim 185, 122]; Zahnersatz kann mit überproportional hohen Frequenzintensitäten bei Zischlauten einhergehen [ArchfKrim 176, 69].

Erkrankungen der Atemwege oder der Stirn- und Nebenhöhlen können zu zeitlich begrenzten sprachlichen Veränderungen führen [ArchfKrim 176, 69].

Beispiel: Erkältungskrankheiten können mit abgesenkten Stimmbandgrundfrequenzen und einer Verbreiterung der Formantenbreiten, Polypen oder durch starken Schnupfen mit einer Nasalierung nichtnasaler Vokale einhergehen [ArchfKrim 176, 69].

Die Zuordnung zu bestimmten pathologischen Erscheinungsgruppen basiert wesentlich auf der Erfahrung des Sprachwissenschaftlers und sollte im Einzelfall sinnvollerweise unter Beteiligung eines Fachmediziners erfolgen [ArchfKrim 176, 69].

2/2.2 Merkmalvergleich

Das Identifizieren von Stimmen erfolgt anhand eines Vergleichs personenspezifischer Merkmale verschiedener Materialproben („minutiöse Gegenüberstellung“ von jeweils zwei Stimmanalysen [NStZ 1989, 404]).

2/2.2.1 Methoden

Da sich die zur Identifizierung von Stimmen verwendeten Merkmale hinsichtlich „des Grades ihrer Objektivierbarkeit“ [NStZ 1989, 404] unterscheiden, kommen für die vergleichende Merkmalanalyse verschiedene Methoden in Betracht.

2/2.2.1.1 Auditive Methoden

Bei den auditiven Methoden werden die für den Merkmalvergleich relevanten Parameter allein mit Hilfe der (akustischen) Wahrnehmung ermittelt und klassifiziert.

Beispiel: Feststellung von Dialekten, Idiolekten, Häsitationsphänomenen, pathologischen Eigenarten, Emotionen [Kriminalistik 1981, 342].

„Analyseinstrument“ ist hier der Mensch mit einem Perzeptionssystem, das durch lebenslange Erfahrung im Bereich der menschlichen Kommunikation verfeinert ist; ein geschulter Fachmann wie etwa ein Phonetiker kann Unterschiede in der Sprechweise feststellen und mit Hilfe eines speziellen Alphabetsystems sehr genau dokumentieren („phonetische Transkription“) [Kriminalistik 1981, 342].

Grundlage aller weiteren Untersuchungen ist die exakte lautliche Umschrift der vorhandenen Sprachproben („segmentelle Transkription des Sprachsignals“ [Kriminalistik 1985, 121]) anhand eines international genormten Transkriptionssystems; erst diese sorgfältig erstellte Transkription ermöglicht die zentrale Aufgabe jedes Stimmvergleichs, nämlich die zuverlässige Ermittlung und Bewertung personenspezifischer Merkmale [Kriminalistik 1985, 122].

Die „Höranalyse“ umfaßt die Untersuchung linguistischer, phonetischer und akustischer, psychischer und pathologischer sowie sozialer und personendiagnostischer Merkmale [Burghard 143]. Wesentlicher Teil der Höranalyse ist der auditive Vergleich analysierter Merkmale zwischen akustischem Spurenmaterial und akustischem Vergleichsmaterial; dabei versucht der Sachverständige im Rahmen des „funktionellen Zuhörens“ die auditiv ermittelten Stimm- und Personenmerkmale aus den verschiedenen Stimmproben zu imitieren, um auf die Bildungsweise dieser Merkmale schließen zu können [Burghard 143].

Die Ergebnisse rein auditiver Verfahren hängen ausschließlich von der persönlichen Sachkunde und Erfahrung des Sachverständigen ab und werden als „entsprechend subjektiv“ bewertet [NSStZ 1989, 404] („subjektive, oder besser: subjektgebundene Methode“ [ArchfKrim 194, 133]). Voraussetzung ist ein durch jahrelanges gründliches Wahrnehmungstraining geschultes Ohr [NSStZ 1989, 403].

2/2.2.1.2 Apparative Methoden

Im Rahmen der „akustischen Meßanalyse“ werden Merkmale der Stimme und der Sprechweise erfaßt, dargestellt, verglichen und bewertet; dabei werden in der Höranalyse erkannte Merkmale meßtechnisch dargestellt und darüber hinaus Merkmale erfaßt, die höranalytisch nicht eingeschätzt werden können [Burghard 8].

Zum Einsatz kommen analoge und/oder rechnergestützte Analysetechniken; während sich Untersuchungen mittels analoger Analysetechnik weitestgehend auf die Darstellung von Merkmalen beschränken, können mit rechnergestützten Verfahren die Analyseergebnisse auch noch statistisch aufbereitet und bewertet werden [Burghard 8].

Analyseverfahren sind unter anderem die Sonographie, die Grundstimmenanalyse und die digitale Signalanalyse [Burghard 8].

Die Schallspektrographie (Sonographie, „Visible Speech“) dient der Darstellung von Sprachschallaufzeichnungen in Abhängigkeit von der Zeit (visueller Vergleich von Kurzzeitspektrogrammen [NSStZ 1989, 404]) und wurde unter dem „irreführenden“ Ausdruck „voiceprint“ (Stimmabdruck) vor allem in den sechziger und siebziger Jahren in den Vereinigten Staaten und teilweise auch in der Bundesrepublik „als (alleinige) Erkenntnisquelle“ für den forensischen Stimmvergleich benutzt [Burghard 265]. Nach massiven Protesten der Fachwelt und einer Reihe von Fehlgutachten in den Vereinigten Staaten wird die Technik „mit der erwähnten Zielsetzung“ in den Vereinigten Staaten kaum noch und in Deutschland „nicht mehr“ angewendet [Burghard 265]. Da sich das Verfahren „jedoch nach wie vor hervorragend zur Darstellung und Dokumentation auditiv-phonetisch ermittelter sprecherspezifischer Merkmale“ eignet, verfügt das Bundeskriminalamt über ein rechnergestütztes System „zur Herstellung hochauflösender Spektrogramme in Falschfarben-Darstellung mit beliebig wählbaren Farbskalen“ [Burghard 265].

Die Grundstimmenanalyse dient der Erfassung der Meßwerte der Grundtonhöhe sowie der Berechnung von statistischen Werten (wie etwa Mittelwert, Minima, Maxima, Standardabweichung, Symmetriekoeffizient); die Untersuchungen werden

mit analogen und digitalen Verfahren durchgeführt [Burghard 134].

Die digitale Signalanalyse ermöglicht auf der Grundlage bekannter mathematischer Verfahren der Spektralanalyse (wie etwa der Fouriertransformation) komplexe physikalisch-akustische Analyseabläufe einschließlich Datengewinnung und Vergleich ausgewählter Analysemerkmale (wie etwa die Grundstimme oder der als Formant bezeichnete stärkste Teilton eines Vokals, der aus der Resonanz in der Mundhöhle entsteht); die jeweiligen Analysedaten können mit Hilfe von Datensammlungen statistisch aufbereitet und zur Untersuchung der Signifikanz einzelner personenspezifischer Merkmale verwendet werden [Burghard 75].

Bei der Automatisierung des Stimmvergleichs werden rein individualtypische Merkmale erfaßt und kompensiert („Sprecher-Normalisierung“), so daß die zur Dekodierung der sprachlichen Nachricht relevanten und nur diese Merkmale hervortreten [NStZ 1989, 401]. Eine vollautomatische Durchführung forensischer Stimmvergleiche („wie eine Atemalkoholanalyse“) ist jedoch „selbst bei optimistischer Betrachtung der Materie als Utopie anzusehen“, da selbst bei der „weitgehend“ automatisierten Messung der durchschnittlichen Grundfrequenz [NStZ 1989, 404] in Abhängigkeit zu den akustischen Gegebenheiten des Sprachmaterials noch immer bestimmte Randbedingungen vom Sachverständigen gesetzt oder kontrolliert werden [NStZ 1989, 405].

Beispiel: Mathematische Fenster- und Filterfunktionen sowie insbesondere die Schwelle zwischen Sprache („Nutzsignal“) und Hintergrundgeräusch („Rauschen“) [NStZ 1989, 405].

2/2.2.1.3 Kombinierte Methoden

Angesichts der Tatsache, daß für die forensische Stimmvergleichsanalyse kein Verfahren in Sicht ist, das einen zur statistischen Berechnung geeigneten Vektor aus personentypischen Merkmalen enthält [Künzel 822], wurde am Bundeskriminalamt als „Symbiose traditioneller ‚geisteswissenschaftlicher‘ und moderner, rechnergestützter, objektivierender – wohlgerneht nicht objektiver! – physikalisch und mathematisch-statistischer

Techniken“ [ArchfKrim 194, 134] ein phonetisch-instrumentelles Kombinationsverfahren entwickelt [NStZ 1989, 402].

Das Hauptmerkmal dieses Verfahrens besteht darin, „daß jeder untersuchte Parameter sowohl durch die Wahrnehmung eines Experten als auch durch die Objektivierung mittels moderner Signalanalyseverfahren in ein Gutachten eingebracht wird“ [NStZ 1989, 404].

Beispiel: Die Bestimmung der durchschnittlichen Grundfrequenz der Stimme ist weitgehend automatisiert und wird mittels einer Hintergrundstatistik ausgewertet; Dialektfärbungen der Sprache sind demgegenüber zunächst ausschließlich subjektiv zu ermitteln, können aber mit Hilfe technischer Verfahren „objektiviert“ („von der Person des Untersuchers unabhängig gemacht“) werden [NStZ 1989, 404].

Bei den apparativen Verfahren handelt es sich um Standardverfahren aus den Bereichen Phonetik, Physik, Medizin und Nachrichtentechnik, die zum Teil auf die speziellen forensischen Randbedingungen angepaßt wurden; neben verschiedenen Formen der Spektralanalyse wird auch eine spezielle Form der Oszillographie benutzt, die eine extreme Auflösung des Sprachsignals im Zeitbereich und dementsprechend exakte Aussagen über temporale Merkmale gestattet, wie etwa die Dauer von Einzellauten oder Lautfragmenten [NStZ 1989, 405].

Ein „Computersystem zur Analyse und Manipulation von Sprache“ ermöglicht unter anderem in Echtzeit die Herstellung farbiger Spektrogramme auf einem hochauflösenden Bildschirm; eine Videokopiereinrichtung ermöglicht die sofortige Herstellung von Abbildungen zur Aufnahme in das Stimmvergleichsgutachten [NStZ 1989, 405].

2/2.2.1.4 Methoden technischer Aufbereitung

Zum Vergleich personenspezifischer Merkmale des lautsprachlichen Verhaltens ist häufig vorab eine technische Bearbeitung der Stimmaufzeichnung erforderlich („Sprachverbesserung“) [NStZ 1989, 401]. Dabei werden gestörte Sprachsignale „nachrichtentechnisch“ aufbereitet, um die „Anhörbarkeit“ für Laien (insbesondere Polizeibehörden und Verfahrensbeteiligte) zu erleichtern.

tern und die „Erkennung“ des übermittelten sprachlichen Inhalts zu verbessern [NStZ 1989, 401].

Hilfsmittel für die „Kompensation oder Abschwächung“ von Störungen aller Art (wie etwa Brummtöne, Rauschen oder Laufgeräusche) sind elektronische Signalverarbeitungstechniken, insbesondere Filterverfahren („Bandpaß, Hochpaß, Tiefpaß, adaptive Filterung“ [Burghard 326]), sowie speziell dafür ausgelegte Rechnersysteme [NStZ 1989, 401]. Der kategorielle Schritt von der absoluten Unverständlichkeit zur völligen Verständlichkeit einer Aufzeichnung ist jedoch in aller Regel auch mit Hilfe modernster rechnergestützter Technik nicht möglich; in diesem Zusammenhang wird auf die im forensischen Bereich „nicht selten“ zu beobachtende zu geringe Aussteuerung von Aufzeichnungen hingewiesen [Burghard 326].

In einigen Fällen ist der Identitätsnachweis durch Stimmvergleich erst im Anschluß an eine derartige nachrichtentechnische Aufbereitung gestörter Sprachaufzeichnungen möglich [NStZ 1989, 401].

2/2.2.2 Dauer der Untersuchung

Je nach Komplexität des Untersuchungsauftrags, dem Vorliegen einer Verstellung und der Menge des Untersuchungsmaterials wird für eine Stimmanalyse ein Zeitaufwand von 30 Minuten bis zu einigen Stunden veranschlagt [NStZ 1989, 404].

2/3 Befund

Die Ergebnisse der vergleichenden Merkmalanalyse bilden den Befund, der sich folglich aus mindestens zwei Stimmanalysen (Spuren- und Vergleichsmaterial) zusammensetzt (die jeweils durch eine Folge von Einzelmerkmalen beschrieben werden [ArchfKrim 176, 70]).

Bei der Erstellung von Stimmvergleichsgutachten hat sich folgende Vorgehensweise bewährt [Kriminalistik 1985, 121]:

- exakte phonetische Umschrift der einzelnen Sprachlaute („segmentelle Transkription des Sprachsignals“) anhand des international genormten Transkriptionssystems;
- zusätzliche Notation von Akzent und Intonation (Satzmelodie);
- Darstellung des daraus gewonnenen Parameter mit Hilfe von Demonstrationsbändern, Kurzzeitspektren, Langzeitspektren, Oszillogrammen oder Autokorrelationsfunktionen (AKF).

Bei dem Vergleich zweier Tonbandaufnahmen ist zu berücksichtigen, daß die jeweiligen Randbedingungen (wie Telefoncharakteristik, Sprechsituation, psychisch-emotionale Verfassung) bei beiden Aufnahmen unterschiedlich und die auditiv wahrgenommenen beziehungsweise apparativ gemessenen Merkmale dadurch verändert sind; die jeweiligen Ergebnisse sind um so präziser, je ähnlicher die Randbedingungen der zu vergleichenden Aufnahmen sind und je mehr qualitativ hochwertige Sprachproben zur Verfügung stehen [ArchfKrim 176, 74]. Bewertungskriterien für den Einfluß von Randbedingungen auf lautsprachliche Merkmale werden in aufwendigen Experimenten statistisch untersucht [ArchfKrim 176, 74] (Erfahrungsbasis!).

Im Rahmen der „Sprecherdiagnostik“ können allein aufgrund einer „Einzelstimmenanalyse“ [ArchfKrim 176, 68] auch Eigenschaften und Merkmale klassifiziert werden, welche die Persönlichkeit der sprechenden Person charakterisieren; diese Merkmale betreffen das psychologische und soziologische Verhalten sowie stimm- und sprachpathologische Besonderheiten und mögliche Stimm- und Sprechverstellungen und werden häufig zur Ermittlung Tatverdächtiger verwendet [Burghard 297]. Aufgrund der Analyse möglichst vieler personenspezifischer Merk-

male sind in der Regel Aussagen zu folgenden Bereichen möglich: Geschlecht, ungefähres Alter, Bildungsniveau, verbales Ausdrucksvermögen, sprachliche und nichtsprachliche akustisch feststellbare Ausdrucksgewohnheiten, Berufszugehörigkeit; ein Zusammenhang zwischen akustisch manifesten Merkmalen der Stimme und bestimmten somatischen Merkmalen wie Größe oder Gewicht wird aus heutiger Sicht klar verneint [NStZ 1989, 404]. Die Sprecherdiagnostik durch Einzelstimmanalyse ist allerdings mit sehr vielen Unsicherheiten verbunden, da sie sich einer wissenschaftlichen Systematik verschließt; ihre Ergebnisse und Interpretation sind weitgehend von der Qualifikation und der Erfahrung des Sachverständigen abhängig [ArchfKrim 176, 68].

Bei der Befundbewertung werden die im Rahmen vergleichender Merkmalanalyse ermittelten Übereinstimmungen und Unterschiede im Hinblick auf ihre sprecherindividualisierende Bedeutung gewichtet [Burghard 298] und in einem „gewichtenden Vergleich“ zu einer Gesamtaussage zusammengefaßt [ArchfKrim 176,70].

Dabei wird grundsätzlich davon ausgegangen, daß Stimme und Sprache infolge ihrer außerordentlichen Komplexität in demselben Maße einmalig sind wie die bekannten äußeren Körpermerkmale [NStZ 1989, 405].

Einschränkend wird allerdings darauf hingewiesen, daß auch im Bereich des lautsprachlichen Verhaltens mehr oder weniger starke Ähnlichkeiten zwischen verschiedenen Individuen zu finden [NStZ 1989, 405] und bestimmte Merkmale der Stimme, Sprache und Sprechweise eines Individuums nicht auf Lebenszeit unveränderlich und einer großen Anzahl linguistischer, situationsbedingter und anderer Variablen [Burghard 298] und Unwägbarkeiten [Kriminalistik 1985, 126] unterworfen sind.

Beispiel: Die Stimme des Babys unterscheidet sich von der des Kindes, nach dem Stimmbruch klingt sie anders als vorher und erfährt weitere Veränderungen über das Erwachsenenalter bis ins Greisenalter [ArchfKrim 176, 66].

Neben Langzeiteffekten treten kurzzeitige und situationsbedingte Veränderungen auf sowie Beeinflussungen durch Emotion und Streß [ArchfKrim 176, 66].

Ferner ist auch die Tatsache zu beachten, daß wegen der außerordentlichen Komplexität der dem Sprechprozeß zugrundeliegenden Bewegungen der Artikulationsorgane und deren neuromuskuläre Steuerung „streng physikalisch betrachtet“ kein Mensch in der Lage ist, eine Äußerung (und sei es nur ein einziger Vokal) zweimal exakt gleich zu reproduzieren [NStZ 1989, 403].

Beispiel: Wird ein in kurzen Zeitabständen von derselben Person fünfzehnmal hintereinander gesprochener Satz einer Spektralanalyse unterzogen und in Form der dazugehörigen Mittelwertspektralkurven ausgedruckt, erhält man statt einer einzigen

Kurve fünfzehn im Detail voneinander abweichende Kurven [ArchfKrim 176, 66].

Andererseits können sprachliche Äußerungen verschiedener Personen dasselbe Frequenzspektrum liefern [ArchfKrim 157, 154].

„Unter diesen Gesichtspunkten ist leicht einzusehen, daß es so etwas wie einen 'Stimmabdruck' mit Reproduktionsqualitäten eines Fingerabdruckes gar nicht geben kann“ [ArchfKrim 176, 66].

Aufgrund der besonderen Eigenart des forensisch bedeutsamen Materials, insbesondere durch den Gebrauch des Telefons, verbleibt zudem im Einzelfall aus der an sich reichhaltigen Anzahl lautsprachlicher Merkmale nur eine jeweils unterschiedlich große Menge auswertbarer Merkmale, deren Zusammensetzung von einem Fall zum anderen danach variiert, welche Bereiche durch die spezifischen Bedingungen des Materials betroffen sind [NSTz 1989, 405]; für den kriminaltechnischen Identitätsnachweis sind daher neben der technischen Qualität und zeitlichen Dauer der Tonaufzeichnung insbesondere „die Anzahl, Struktur und Stabilität der vergleichbaren sprecherindividualisierenden Merkmale“ von ausschlaggebender Bedeutung [Burghard 8].

Grundlegend für den Identitätsnachweis ist die genaue Kenntnis personenspezifischer Unterschiede des lautsprachlichen Verhaltens, und zwar sowohl hinsichtlich verschiedener Personen als auch solcher, die für ein und dieselbe Person gelten [NSTz 1989, 402]. Zur Isolierung der Merkmale, bei denen die durch Anlage (Erbgut, Geschlecht, Alter) und Umwelt (Region, Schichtzugehörigkeit, Bildung) bedingten personenspezifischen Unterschiede zwischen verschiedenen Personen möglichst groß und Unterschiede im lautsprachlichen Verhalten ein und derselben Person möglichst klein sind, „waren und sind zahlreiche empirische Untersuchungen erforderlich“ [NSTz 1989, 403]; die forensische Tauglichkeit einzelner Merkmale wird anhand empirischer Testreihen ermittelt, wobei insbesondere auch die „Robustheit“ des Merkmals gegenüber Verstellung und anderen typischen Randbedingungen forensischer Fragestellungen zu berücksichtigen ist [NSTz 1989, 405].

Im konkreten Fall nimmt die personenspezifische Potenz einzelner Merkmale in dem Maße zu, wie sie von durchschnittlichen Werten abweicht [NSTz 1989, 404].

Beispiel: Werden für eine anonyme Person und eine Vergleichsperson jeweils gleiche Werte der Stimmbandgrundfrequenz ermittelt, kann eine Wahrscheinlichkeit dafür angegeben werden, daß rein zufällig zwei Personen eine so tiefe oder noch tiefere beziehungsweise so hohe oder höhere Stimme haben (der angegebene Wahrscheinlichkeitsgrad beinhaltet allerdings keine Feststellung zur Identität der Stimmen, da die durchschnittliche Stimmfrequenz nur ein Merkmal unter vielen ist) [NSTz 1989, 403].

Bei allen Merkmalen (außer dem Dialekt) können zudem pathologische Ausprägungen auftreten (Stimm-, Sprach- und Sprechstörungen), die eine erhebliche Steigerung ihres personenspezifischen Wertes bewirken [Künzel 826].

Beispiel: Bei Vorliegen einer nicht auf Verstellung beruhende fehlerhaften Aussprache der 's'-Laute (Sigmatismus) kann der Ausschluß von 95% aller (erwachsenen) Personen erfolgen; je nach Ausprägung des Sigmatismus als Fehlartikulation etwa der Zungenspitze (interdental) oder des Zungenrandes (lateral) sind noch weitergehende Eingrenzungen möglich [NSTz 1989, 404].

Die Ergebnisse der forensischen Stimmvergleichsanalyse basieren demnach neben den Besonderheiten des Einzelfalles und der Kompetenz des Sachverständigen ganz entscheidend auf Bestand und Zuverlässigkeit wissenschaftlich abgesicherter Fakten und Zahlen hinsichtlich der untersuchten Merkmale [Kriminalistik 1985, 125].

Beispiel: Statistische Erkenntnisse über die Verteilung eines bestimmten Sprachfehlers in der Bevölkerung des Landes, über die durchschnittliche Grundfrequenz einer Stimme bei erwachsenen oder heranwachsenden männlichen oder weiblichen Personen, über die zahlenmäßige Verbreitung eines bestimmten Dialektes oder Subdialektes sowie Statistiken über bestimmte Merkmale wie Verzögerungsphänomene, Sprechtempo oder Satzmelodie [Kriminalistik 1985, 125].

Da jedoch weder die gesamte Anzahl noch die Interdependenz der verschiedenen personenspezifischen Merkmale bekannt sind, scheidet die Beurteilung der Identitätsfrage anhand einer linearen Verhältnisskala von 0 bis 1 bzw. 100% aus [NSTz 1989, 405].

Für die abschließende Bewertung des Untersuchungsbefundes wird folgende Skalierung vorgeschlagen [NStZ 1989,405]:

Identität bzw. Nichtidentität

- kann nicht beurteilt werden;
- ist mit gewisser Wahrscheinlichkeit anzunehmen;
- ist mit großer Wahrscheinlichkeit anzunehmen;
- ist mit sehr großer Wahrscheinlichkeit anzunehmen;
- ist ohne jeden vernünftigen Zweifel anzunehmen.

Die letzte Wertung bleibt allerdings Fällen vorbehalten, in denen große Mengen unter hochwertigen akustischen Bedingungen aufgezeichneten Materials vorliegen und die sprechenden Personen besonders markante (pathologische) Merkmale aufweisen; Telefonaufzeichnungen kommen für diese Kategorie nicht in Betracht [NStZ 1989, 405].

2/5 Fehlerquellen

Im Folgenden werden einige Fehlerquellen der vergleichenden Merkmalanalyse behandelt (siehe dazu allgemein „Identifizieren und Wiedererkennen“, „Sachverständigengutachten“ und „Spurenuntersuchung und Auswertung“; zu Fragen der Verteidigung im Zusammenhang mit kriminaltechnischen Befunden siehe „Spurenuntersuchung und Auswertung“).

2/5.1 Material

2/5.1.1 Zeitlicher Umfang der Tonaufzeichnung

In der forensischen Praxis sind Tonaufzeichnungen von nur 5 Sekunden Dauer keine Seltenheit; eine Dauer von 60 Sekunden ist bereits als überdurchschnittlich anzusehen [NStZ 1989, 402].

Die Anzahl der in einem Satz möglichen Wörter ist in forensischen Spurenproben im Prinzip unbegrenzt [NStZ 1989, 402].

2/5.1.2 Telefonaufzeichnung

In mehr als 95% der forensisch bedeutsamen Fälle wird zur Tatusführung das Telefon benutzt, was zu verschiedenen Beeinträchtigungen der akustischen Qualität des Sprachsignals führt [NStZ 1989, 402].

Beispiel: Einschränkungen des Lautstärkebereichs („Dynamikbereichs“) auf bestenfalls ca. 30 Dezibel; Verzerrungen verschiedener Stärke und mathematischer Komplexität; Brummstörungen aus dem Lichtnetz [NStZ 1989, 402]; Raumakustik; Frequenzgangfehler durch den meist behelfsmäßig ausgeführten Anschluß des zur Aufzeichnung benutzten Tonbandgerätes [ArchfKrim 157, 154].

Durch die Beschränkung des Frequenzbereiches auf 300 bis 3400 Hz (Hertz) werden tiefere und/oder höherer Frequenzanteile unterdrückt, so daß etwa „potente“ personenspezifische Merkmale hochfrequenter Laute nicht oder nicht sicher festgestellt werden können [NStZ 1989, 402].

Beispiel: Als Folge der Frequenzbeschnidung des Übertragungskanalns können die Laute 's' und 'f' im allgemeinen nicht unterschieden werden, da beide akustisch durch die Filterung weitgehend angeglichen werden; für die sprachliche Verständigung bereitet dies normalerweise keine Schwierigkeiten, da aufgrund des sprachlichen Zusammenhangs etwa zwischen „Lauf“ und „Laus“ unterschieden werden kann, während ohne inhaltliche Information, etwa bei Eigennamen, Buchstabierhilfen erforderlich sind: „... ich meine Herrn Hufmann mit 'f' wie Friedrich, nicht 's' wie Siegfried“ [ArchfKrim 194, 132].

Personenspezifische Merkmale werden durch Telefonübertragung „maskiert“ [Künzel 838]; so gilt etwa der Merkmalbereich Stimmqualität für die forensische Anwendung als nahezu wertlos, weil er in ausgeprägter und überwiegend nicht kompensierbarer Weise von den Einflüssen der Telefonübertragung beeinträchtigt wird [Künzel 824].

Beeinträchtigungen der Tonqualität werden in vielen Fällen durch unsachgemäße Aufzeichnung von Telefongesprächen erzeugt oder verstärkt [NStZ 1989, 402] (Einfluß der Ermittlungsbeamten auf den Beweiswert sachlicher Beweismittel!).

Allein der langandauernde tägliche Umgang mit telefonübertragenen Sprachsignalen vermittelt die zur Stimmvergleichsanalyse notwendige Erfahrung [Kriminalistik 1985, 122].

2/5.1.3 Akustische Besonderheiten

2/5.1.3.1 Verstellung

Zur Verschleierung der Identität wird in ca. 15% der forensisch bedeutsamen Fälle das lautsprachliche Verhalten aktiv verstellt; aus der Praxis sind zahlreiche „die Erstellung eines Gutachtens in unterschiedlichem Maße beeinträchtigende“ Varianten und Kombinationen von Manipulationen der Stimme, Sprache und Sprechweise bekannt [NStZ 1989, 401] (Erfahrungsbasis!).

Veränderungen der Stimme und Sprechweise sind in gewissen Grenzen durch bewußte Beeinflussung des lautsprachlichen Verhaltens möglich, wobei jedoch die Verstellbarkeit einzelner Merkmale recht unterschiedlich ausfällt: als leicht verstellbar

gelten Stilschicht (Wortwahl, Satzbau, Ausdrucksweise), Stilfärbung und Dialekt (Ganzheitsmerkmale), Artikulation, Akzentuierung, Sprechfluß und Sprechtempo (Einzelmerkmale); als schwer oder überhaupt nicht verstellbar gelten Stimm- und Sprechstörungen, Sprechdynamik, Tonhöhe, Lautstärke, Klangfarbe, Sprechrhythmus und Stimmfülle [Burghard 199]. Versuchsergebnisse deuten darauf hin, daß sich selbst sprachlich geübte Personen maximal nur auf vier bis fünf Merkmale in einer Verstellungsvariante gleichzeitig konzentrieren und dabei diese Variante nur relativ kurze Zeit konsequent durchhalten können; erkennbar wird dies an der Tendenz, mit der Zeit immer wieder in die natürliche Stimme und Sprechweise zurückzufallen [Burghard 299]. Verstellungen sind auch an ihrer Unnatürlichkeit und einem erhöhten Spannungszustand erkennbar [Burghard 299].

Eine spezielle Art der Verstellung ist die Imitation der Stimme oder Sprechweise einer anderen Person; selbst bei äußerlich verblüffender Ähnlichkeit bieten jedoch vor allem die akustische Meßanalyse sowie der direkte höranalytische Vergleich mit dem Originalsprecher Möglichkeiten einer eindeutigen Differenzierung [Burghard 299].

Zu beachten sind auch mechanische Manipulationen wie Nase-zuhalten oder Blockieren der Zunge [Burghard 305].

Bei bestimmten Arten der Stimmverstellung (etwa dem Flüstern) werden geschlechtsdifferenzierende Merkmale wie Stimmfrequenz und Stimmklang maskiert [NStZ 1989, 404].

2/5.1.3.2 Natürliche Stimmprofilveränderung

Eine natürliche Veränderung des Stimmprofils kann etwa durch Erkältung oder Alkoholmißbrauch eintreten [NStZ 1989, 402].

Für das Phänomen „Heiserkeit“ werden verschiedene Merkmale benutzt, darunter die Unregelmäßigkeit der Stimmtoneerzeugung und das Ausmaß nichtperiodischer Anteile (Rauschenergie) bei der Stimmtoneerzeugung, für welches derzeit wegen der im forensischen Bereich fast ausschließlich gegebenen Telefonübertragung kein zuverlässiges Meßverfahren bekannt ist [ArchfKrim 194, 138].

2/5.1.3.3 Sprechsituation

Zwischen verschiedenen Materialproben können „pragmatische und situative Unterschiede“ bestehen, welche einen Vergleich erschweren und die Aussagekraft bestimmter sprachlicher Merkmale von vornherein einschränken [ArchfKrim 198, 3].

Beispiel: Materialprobe mit laut vorgelesenem Text, sehr langsam und mit deutlicher, zum Teil auch emphatischer Hervorhebung wichtiger Aussageinhalte durch Betonung und Melodieführung (Intonationsverlauf) gegenüber einer Vergleichsprobe, die ausschließlich aus kurzen, spontanen Antworten auf Fragen von Journalisten besteht [ArchfKrim 198, 3].

Randbedingungen, die Inhalt und Form einer sprachlichen Äußerung mitbestimmen, werden sowohl von der Sprechenden als auch von der Zuhörenden Person und den äußeren Umständen gesetzt [Burghard 299].

Beispiel: Charaktereigenschaften, Motivation oder Erregung auf Seiten der Sprechenden Person; Interesse, Aufmerksamkeit oder Gegenrede auf Seiten der Zuhörenden Person; Räumlichkeit oder Geräusche als äußere Einflüsse [Burghard 299].

Die Berücksichtigung sämtlicher Randbedingungen ist für den forensischen Stimmvergleich von entscheidender Bedeutung, da eine größere Anzahl von Merkmalen der Stimme und der Sprechweise mehr oder weniger unmittelbar von der konkreten Sprechsituation abhängen und von ihr beeinflusst werden [Burghard 299].

Beispiel: Weist eine Sprachprobe eine besondere Art von Sigmastismus auf, eine andere dagegen nicht, kann dies ein Anzeichen für Nichtidentität sein; es kann aber auch bei der zweiten Sprachprobe lediglich eine schlechte Telefonverbindung vorgelegen haben (hier könnte eine Langzeit-Spektralanalyse zur Klärung beitragen) [Kriminalistik 1985, 125].

Die Entscheidung, welche Merkmale von Stimme, Sprache und Sprechweise zum Vergleich herangezogen werden können, setzt viel Erfahrung voraus [Kriminalistik 1985, 125].

2/5.2 Methode

2/5.2.1 Technische Bearbeitung des Spurenmaterials

Da Spurenmateriale in vielen Fällen „nachrichtentechnisch“ aufbereitet werden und Vergleichsproben in aller Regel unter anderen akustischen Bedingungen (und möglicherweise mit anderen sprachlichen Inhalten) aufgezeichnet werden, dürfte sich in allen Fällen die grundsätzliche Frage einer Vergleichbarkeit des Untersuchungsmaterials stellen. Dabei ist insbesondere zu berücksichtigen, daß selbst bei weitgehend automatisierten Verfahren „in Abhängigkeit der akustischen Gegebenheiten des Spurenmaterials“ bestimmte Randbedingungen „noch immer“ durch den Sachverständigen selbst gesetzt und kontrolliert werden [NStZ 1989, 405].

Beispiel: Mathematische Fenster- und Filterfunktionen sowie insbesondere die Schwelle zwischen Sprache („Nutzsprache“) und Hintergrundgeräusch („Rauschen“) [NStZ 1989, 405].

Die zur technischen Qualitätsverbesserung von Tonaufzeichnungen eingesetzten Verfahren „arbeiten zwar computergestützt, aber damit keinesfalls vollautomatisch, sondern vielmehr interaktiv“; der Benutzer kann eine Vielzahl von Variablen beeinflussen, „die beträchtlichen phonetischen und signalanalytischen Sachverstand voraussetzen“ und die exakte Anpassung der „Voreinstellung des Rechners“ auf die Erfordernisse des Einzelfalles gewährleisten sollen [NStZ 1996, 178]. Damit soll vor allem vermieden werden, „daß durch die technische Bearbeitung Artefakte entstehen, deren schlimmstmögliche Konsequenz eine falsche Textverschriftung wäre“ [NStZ 1996, 178] (und damit die „Sprachverbesserung“ zur „Sprachverwässerung“ entartet).

Die Präsentation und Erläuterung mehrfarbiger Spektralbilder hilft über dieses grundsätzliche Problem nicht hinweg, da diese Hilfsmittel in aller Regel den subjektiv erhobenen Untersuchungsbefund nur visuell darstellen („objektivieren“), darüber hinaus aber keine Erkenntnisse über die (vom Sachverständigen vorab bestimmten) Umstände und Voraussetzungen dieser „Objektivierung“ vermitteln.

Eine detaillierte Abklärung der Frage, ob und inwieweit eine eventuelle technische Bearbeitung des Untersuchungsmaterials

die akustische Meßanalyse oder schlimmstenfalls die Textverschriftung beeinflußt, kann daher im Einzelfall durchaus von einigem Interesse sein.

Beispiel: Ist die Tonaufzeichnung infolge nicht hinreichender Erdung beziehungsweise Abschirmung durch einen Oberton der Netzfrequenz (in Europa 50 Hertz) oder einen einfachen monofrequenten Pfeifton beeinträchtigt und wird diese Störung mittels „einer steilen Kerbfilterung“ beseitigt, geht zwangsläufig auch die sprachliche Information in dem betroffenen Frequenzbereich verloren; im Spektrum entsteht ein „Frequenzloch“ [ArchfKrim 194, 134].

In manchen Fällen kann es vorteilhafter oder sogar unabdingbar sein, eine Störung nicht vollständig zu unterdrücken sondern lediglich abzuschwächen, um die Natürlichkeit des Sprachsignals und damit den charakteristischen Klang der Stimme möglichst wenig zu beeinträchtigen; um hier den „Königsweg“ zu finden, sind sprachwissenschaftliche Kenntnisse erforderlich [ArchfKrim 194, 136]. In Ausnahmefällen kann ein und derselbe Text durch eine technische Aufbereitung überhaupt nicht und unter anderen Bedingungen zu 100 Prozent verständlich sein [ArchfKrim 194, 136].

2/5.2.2 Korrekte Meßvoraussetzungen

Vor der Messung der Stimmbandgrundfrequenz aus Tonbandmaterial ist festzustellen, mit welcher exakten Aufnahme- und Wiedergabegeschwindigkeit gearbeitet wurde [Kriminalistik 1981, 345].

Beispiel: Läuft das zur Aufnahme benutzte Gerät etwas langsamer als das zur Wiedergabe benutzte Gerät, erscheint die Grundfrequenz um den entsprechenden Betrag zu hoch oder umgekehrt [Kriminalistik 1981, 345].

2/5.2.3 „Subjektiver Faktor“

Überall dort, wo ein „menschlicher Untersucher“ beteiligt ist, kann „ein minimales subjektives Moment“ nicht vermieden werden; dies gilt für jede Art forensischer Untersuchung, sei es auf

dem Gebiet der Chemie, Physik, Medizin (Psychiatrie), Psychologie und insbesondere für die mit speziellen Techniken der Merkmalextraktion und Merkmalsbewertung durchgeführte phonetische Sprechererkennung [Kriminalistik 1981, 126].

2/5.2.4 Visualisierung

Zur „Substantiierung der Meinungsbildung des Gerichts“ oder als Grundlage einer Diskussion mit „eventuellen“ Ko- oder Gegengutachtern werden in aller Regel visuelle Darstellungen verwendet, vor allem bezüglich „ausschließlich subjektiv“ ermittelter und durch technische Verfahren „objektivierter“ Merkmale [NStZ 1989, 404].

Die forensische Gutachterpraxis hat zu der Erkenntnis geführt, daß „dem Laien“ infolge der vorübergehenden (transitorischen) Natur der gesprochenen (im Gegensatz zur schriftlich fixierten) Sprache zwar bis zu einem gewissen Grad eine Vielzahl lautlicher Charakteristika durch Tonbänder demonstriert und verdeutlicht werden können, letztendlich überzeugend aber erst die Vorlage eines „beliebig lange zu betrachtenden Analyseprotokolls“ wirke, das etwa als Computerspektrogramm den betreffenden Sachverhalt graphisch darstellt, „obwohl durch diese Verlagerung der Wahrnehmung von der auditiven auf die visuelle Ebene die im Sprachsignal enthaltene akustische Information auf einen Bruchteil reduziert wird“ [NStZ 1989, 404] (eine Videodemonstration in Bild und Ton wäre daher möglicherweise weitaus sachdienlicher).

Neben dieser Verlagerung der Wahrnehmungsart suggeriert der Begriff „Analyseprotokoll“ im Zusammenhang mit computerfertigtem Anschauungsmaterial die Vorstellung einer vollautomatischen Durchführung der Stimmvergleichsanalyse („wie eine Atemalkoholanalyse“ [NStZ 1989, 405]), die jedoch bei den von den Polizeibehörden verwendeten Kombinationsverfahren überhaupt nur für einen beschränkten Merkmalbereich in Betracht kommt und dort auch nur als „halbautomatisch“ bezeichnet werden kann.

Weitaus informativer als spektakuläres Bildmaterial wäre möglicherweise in vielen Fällen eine „minutiöse“ [NStZ 1989, 404] Darstellung der gesamten Erhebung und Bearbeitung des Unter-

suchungsmaterials einschließlich einer kritischen Beleuchtung etwaiger zugrundeliegenden Subjektivismen und sonstiger Fehlerquellen.

2/5.3 Erfahrungswissen

Unter Umständen ist das dem Stimmvergleich zugrundeliegende empirische Material (Erfahrungsbasis) im Hinblick auf Art, Umfang, tatsächlichen Voraussetzungen und Aussagekraft (Anwendungsbereich) einer eingehenden fallbezogenen Prüfung und Kontrolle zu unterziehen (siehe dazu „Erfahrungssätze“).

2/5.4 Bewertung

Bei dem forensisch bedeutsamen Vergleich der Stimme einer anonymen Person (Straftäter) und einer Vergleichsperson (Tatverdächtiger) ist stets zu beachten, daß die Menge der infragekommenden Personen außerordentlich groß sein kann [NStZ 1989, 402].

Beispiel: Sämtliche männliche Personen einer bestimmten Altersklasse mit Deutsch als Muttersprache [NStZ 1989, 402].

Zudem kann nicht davon ausgegangen werden, daß sich die gesuchte Person unter den drei oder auch dreißig Personen befindet, deren Sprachproben zum Vergleich untersucht wurden [NStZ 1989, 402].

2/6 Resümee

Sachverständigengutachten zum forensischen Stimmvergleich werden in Deutschland „fast ausschließlich“ [GA 1988, 215] vom Bundeskriminalamt sowie den Landeskriminalämtern Bayern, Nordrhein-Westfalen und Brandenburg [Burghard 297] als Behördengutachten im Sinne des § 256 StPO erstellt und vor Gericht vertreten“ [GA 1988, 215]. Polizeiunabhängige „Ko- oder 'Gegen'-gutachter“ [GA 1988, 215] dürften demgegenüber ausgesprochen selten sein.

Das im Bundeskriminalamt entwickelte und von verschiedenen Landeskriminalämtern praktizierte „phonetisch-instrumentelle Kombinationsverfahren“ wurde „in weit über 1000 Kriminalfällen erfolgreich eingesetzt“; entsprechende Gutachten sind, „soweit eine Erstattung in der Hauptverhandlung erforderlich war, ohne Ausnahme von den Gerichten akzeptiert worden“ [NStZ 1989, 405].

Angesichts einer (vermutlich) weitgehend kritiklosen Rezeption von Behördengutachten durch die Tatgerichte sollte die „Erfolgsbilanz“ der forensischen Stimmanalytik nicht darüber hinwegtäuschen, daß die neben den nach wie vor unersetzlichen „auditiven“ Methoden zur „Objektivierung“ subjektiv ermittelter Befunde verwendeten technischen Verfahren mehr oder weniger ausschließlich vom Bundeskriminalamt selbst entwickelt und angewendet werden.

Eine wissenschaftliche Überprüfung dieser für die spezifisch forensischen Bedingungen entwickelten Verfahren durch polizeiunabhängige Sachverständige dürfte in vielerlei Hinsicht noch ausstehen. Da „aus einer Reihe von Gründen .. bestimmte technische Details .. sowie einige als Vorbereitung auf die eigentlichen Untersuchungen der Sprachsignale angewandte technische Verfahren“ überhaupt nicht publiziert werden [ArchfKrim 198, 2], wird eine derartige Überprüfung kriminaltechnischer Erkenntnisse und Methoden im Bereich der forensischen Stimmanalytik sogar gänzlich ausgeschlossen sein.

Soweit auf Seiten der Verteidigung eine Auseinandersetzung mit dem Ergebnis kriminaltechnischer Gutachten für erforderlich gehalten wird (siehe dazu „Spurenuntersuchung und Auswertung“) sollten Stimmvergleichsanalysen trotz (oder gerade we-

gen) ihrer großen Akzeptanz bei den Tatgerichten stets kritisch gewürdigt werden – insbesondere im Hinblick auf die polizeiinterne Entwicklung und Verwendung eines speziellen Computersystems „zur Analyse und Manipulation von Sprache“ [Künzel 834] und aufgrund der Tatsache, „daß ein instrumentell (d.h. vor allem rechnerisch) gut ausgerüstetes und über spezielles sprachwissenschaftliches Know-how verfügendes Tonstudio im Auftrag interessierter Kreise“ durchaus in der Lage ist, „eine Fälschung“ zu produzieren [ArchfKrim 198, 2].

3 Wiedererkennen von Stimmen

„Wiedererkennen“ wird hier im Sinne eines Identitätsnachweises aufgrund subjektiver Erinnerung verstanden (siehe „Identifizieren und Wiedererkennen“).

Subjektive Erinnerung kann als Vorgang anschaulichen Erlebens beschrieben werden, der zwar ebenso wie objektbezogene Analysen „objektive Gewißheit“ intendiert, aber nur „subjektive und relative Sicherheit“ erreicht, da individuelle Äußerungen zur Identität nur schwer zu begründen und ihrer Überprüfbarkeit enge Grenzen gesetzt sind [ArchfKrim 176, 144].

Wer sich aufgrund aktueller Eindrücke an menschliche Stimmen erinnert, ist im Gegensatz zur objektbezogenen Analyse nicht mit seinem analytischen Verstand sondern mit einer „ganzheitlichen Wahrnehmung des Gesamten“ (Anschauung) gefordert, die eher mit dem „Angemutetsein des Bekannten und Vertrauten“ zu tun hat und weniger das Einzigartige als vielmehr das Typische und Charakteristische erfaßt (Gestaltqualität) [ArchfKrim 176, 144].

In der forensischen Praxis kann das Wiedererkennen von Stimmen grundsätzlich in zweierlei Hinsicht von Bedeutung sein: zur Ermittlung von Tatverdächtigen durch die Präsentation von tatrelevanten Stimmspuren gegenüber einem unbestimmten Publikum („akustischer Fahndungsaufruf“) und zum Nachweis der Täterschaft durch die Präsentation akustischer Stimmproben von Tatverdächtigen gegenüber „Ohrenzeugen“ („akustische Gegenüberstellung“ [Bender 195]).

Die folgenden Ausführungen beschränken sich auf die Fallgruppe „akustische Gegenüberstellung“ im Rahmen des Identitätsnachweises.

3/1 Hilfsmittel

Als Hilfsmittel für das Wiedererkennen von Stimmen aufgrund subjektiver Erinnerung dienen aktuelle Sprechproben oder Tonaufzeichnungen.

Für die Präsentation des akustischen Materials gegenüber einem oder mehreren Ohrenzeugen sind grundsätzlich folgende Möglichkeiten in Betracht zu ziehen:

- visuelle Gegenüberstellung mit Sprechproben;
- Videoaufzeichnungen einer visuellen Gegenüberstellung mit Sprechproben;
- akustische Gegenüberstellung unter Ausschluß visueller Wahrnehmungen („akustische Realgegenüberstellung“);
- akustische Gegenüberstellung mittels Tonaufzeichnungen.

Die Arrangements können in neutraler Umgebung („Tatferne“) oder in räumlicher Nähe zum Tatgeschehen erfolgen („Tatnähe“); hinsichtlich der akustischen Wahrnehmungen der Beweisperson ist jedoch stets zwischen den akustischen Wahrnehmungen in der Tatsituation und denen in der Vernehmungssituation („tatfern“ oder „tatnah“) zu unterscheiden.

Zu den Voraussetzungen und zum Ablauf akustischer Gegenüberstellungen siehe unten unter Fehlerquellen (Gegenüberstellung).

3/2 Typisierendes Erleben

Das „innerpsychische Wirkgefüge“, welches der Äußerung einer Beweisperson zur Identität menschlicher Stimmen letztendlich zugrunde liegt, gilt im Gegensatz zum Erkenntnisvorgang der objektbezogenen Analyse als nur unzureichend beschreibbar [ArchfKrim 176,144]. Vorausgesetzt wird in jedem Fall eine erlebnisbegründete Bekanntschaft mit dem Gegenstand des Wiedererkennens, also ein „Anknüpfen an gespeicherte Informationen“ oder ein „Rückgriff auf eine Datenbasis“ sowie eine „wahrnehmungsnah reproduktion des Erinnerungsbildes“ [ArchfKrim 176,144].

Über Art und Funktionsweise des Stimmgedächtnisses scheinen auch in Fachkreisen gewisse Unklarheiten zu bestehen. Es gilt als schwierig, „ein schlüssiges Bild“ der „Stimmenidentifizierung“ zu bekommen: die einschlägigen wissenschaftlichen Erkenntnisse (Erfahrungsbasis!) werden auf ganz unterschiedlichen Forschungsgebieten gewonnen, so daß „die Befunde teilweise uneinheitlich, ohne gegenseitige Bezugnahme, und häufig aufgrund unterschiedlicher Erhebungsmethoden nicht vergleichbar“ sind [Hammersley 120] („zu unterschiedlich sind die bislang eingesetzten Testverfahren und zu heterogen die Fragestellungen“ [GA 1988, 224]). „Streng auf forensische Verhältnisse ausgerichtete Studien sind ausgesprochen selten“ [GA 1988, 224].

Viele Anhaltspunkte sprechen dafür, daß Menschen normalerweise keine Informationen über die akustische Qualität einer Stimme speichern, solange sie nicht ausdrücklich darauf achten oder solange die Besonderheiten der Stimme nicht in Zusammenhang mit dem Inhalt des Gesprochenen stehen [Hammersley 118].

Es wird vermutet, daß akustische Informationen als solche vom Gedächtnis überhaupt nicht oder jedenfalls in einer noch unbekannt Form gespeichert werden, und daß „stimmliche Informationen“ vor der Speicherung „umkodiert“ werden [Bender 193].

Beispiel: Kodierung akustischer Signale als „männliche Stimme“ [Bender 193].

Jedenfalls aber wird keine „Kopie des Originals“ gespeichert sondern zumeist nur die Bedeutung des Gesprochenen, die möglicherweise noch mit dem einen oder anderen zusätzlichen „Informationsbit“ versehen ist [Bender 193].

Beispiel: „Panik“ (Bedeutung) und „männliche Stimme“ (Zusatzinformation).

Stimmliche Informationen werden nur dann im Gedächtnis behalten, wenn sie für den Hörer oder für den Inhalt der Information von Bedeutung sind [Bender 193]. Die Bedeutung kann sowohl einen Bezug zur gegenwärtigen Aktivität des Hörers als auch zu seinen Erwartungen von zukünftigen Ereignissen haben; langfristig wird statt der akustischen Merkmale der Stimme eher die Bedeutung des Gesprochenen behalten [Hammersley 118].

Beispiel: Emotionaler Zustand des Sprechers [Hammersley 118].

Zur Erklärung wird auf die unterschiedlichen Funktionen des Kurzzeit- und Langzeitgedächtnisses verwiesen: das Kurzzeitgedächtnis kodiert die akustischen und visuellen Eigenschaften von Objekten, das Langzeitgedächtnis deren Bedeutung [Bender 193].

Das „Behalten bestimmter Aspekte stimmlicher Informationen“ bedeutet nicht, daß der Hörer „eine exakte akustische Repräsentation der Stimme im Gedächtnis hat“ [Hammersley 119]. Im Hinblick auf den inneren Vorgang des Wiedererkennens wird vermutet, daß die Beweisperson das ihr präsentierte stimmliche Material nicht mit einer „Gedächtnisrepräsentation“ vergleicht sondern zu rekonstruieren versucht, „wie die Stimme des Täters geklungen haben muß“ [Bender 193].

All diese Vermutungen verweisen auf den ganzheitlichen Charakter der Wiedererkennensleistung und stützen die Annahme, daß beim Wiedererkennen menschlicher Stimmen aufgrund subjektiver Erinnerung das für die objektbezogene Analyse typische Vergleichen einzelner Merkmale „gewöhnlich nicht“ [ArchfKrim 176, 144] auftritt.

Das Ergebnis „typisierenden Erlebens“ ist eine Aussage hinsichtlich einer Übereinstimmung oder Verschiedenheit des im Rahmen einer Gegenüberstellung präsentierten akustischen Materials (Stimmen von Tatverdächtigen und/oder Vergleichspersonen) mit den akustischen Wahrnehmungen während der Tatsituation (Täterstimme).

Kann die Beweisperson das präsentierte Material nicht eindeutig zuordnen („Identifizierung“ [Hammersley 115]), kommt als Ergebnis auch in Betracht:

- die Bekundung der Beweisperson, eine Stimme schon einmal gehört zu haben („Bekanntheit“ [Hammersley 115]);
- die Fähigkeit der Beweisperson, eine bekannte Stimme aus mehreren verschiedenen Stimmen herauszuhören („Wiedererkennen“ [Hammersley 115]).

Diese Differenzierung entspricht in etwa der zwischen dem „individualisierenden“ und „gattungsspezifischen“ Wiedererkennen (dazu und zur abweichenden Terminologie siehe „Identifizieren und Wiedererkennen“; die terminologische Differenzierung von Hammersley/Read beruht offenbar auf der Sichtung einschlägiger wissenschaftlicher Studien und dürfte damit in erster Linie in den unterschiedlichen Fragestellungen und Versuchsanordnungen dieser Untersuchungen begründet sein [Hammersley 115]).

3/4

Ergebniskontrolle

Als „ganzheitliches Erleben“ resultiert die Wiedererkennensleistung des Ohrenzeugen nicht nur aus der individuellen Wahrnehmungs- und Erinnerungsfähigkeit sondern ganz entscheidend aus höchstpersönlichen „Konfigurationen“ der Beweisperson, die als solche weitestgehend einer Überprüfung und Kontrolle entzogen sind.

Beispiel: Die Bereitschaft einer Beweisperson, auch bei subjektiver Unsicherheit eine bestimmte Zuordnung vorzunehmen, kann mittelbar durch seine Erziehung und von seinem Selbstbewußtsein beeinflusst sein [Hammersley 128].

Zur Ergebniskontrolle kann ein Sachverständiger zugezogen werden, der sich unter anderem auch zur „Erkennungsstrategie“ der Beweisperson äußern kann [Künzel 839]. Mitunter werden dabei allerdings auch sachverständigerseits grundlegende Differenzierungen nur unzureichend erfaßt.

Beispiel: „Die Untersuchung des Sachverhalts im konkreten Fall wird zweckmäßigerweise von einem Sachverständigen vorzunehmen sein, der darüber hinaus die Erkennungsstrategie einer Person beurteilen kann, denn für das Gericht ist auch von Interesse, ob eine positive Erkennung anhand konkreter, vom Sachverständigen dann in wissenschaftliche Terminologie zu fassender sprachlicher Merkmale begründet wird, oder ob eine nicht differenzierte, ganzheitliche Erkennung vorliegt“ [Künzel 839]; „während die Mehrzahl der Beurteiler eine ganzheitliche Entscheidungsstrategie verfolgt, ist eine Reihe von ihnen in der Lage, zumindest zwei oder drei sprechertypische Merkmale in vorwissenschaftlicher Terminologie auszudrücken“ [GA 1988, 219].

Die sprachliche Begründung eines Wiedererkennens (anhand konkreter begrifflich faßbarer Merkmale) und der subjektive Vorgang des Wiedererkennens selbst (aufgrund ganzheitlichen Erlebens) sind ihrer Art nach durchaus verschiedene psychische Leistungen der Beweisperson. Es ist deshalb zweifelhaft, ob diese unterschiedlichen psychischen Leistungen als solche überhaupt in eine alternative Relation gebracht werden können, da sich als Alternative zur sprachlichen Begründung eines Wiedererkennens anhand konkreter Merkmale zunächst einmal der Umstand aufdrängt, daß die Beweisperson nicht in der Lage ist, zur

Begründung ihrer Identitätsbehauptung bestimmte Kriterien zu benennen. Desweiteren erscheint durchaus fraglich, ob allein das sprachliche Unvermögen der Beweisperson auf eine „nicht differenzierte ganzheitliche Erkennung“ hindeutet, da beim Wiedererkennen das für die objektbezogene Analyse typische Vergleichen bestimmter Merkmale „gewöhnlich“ gerade nicht auftritt. Es ist deshalb durchaus möglich, daß sich die Beweispersonen tatsächlich nicht durch verschiedene „Erkennungsstrategien“ sondern allein durch ihr sprachliches und begriffliches Ausdrucksvermögen unterscheiden; dies wäre jedenfalls angemessen zu problematisieren.

Im übrigen ist zu beachten, daß sich das Wiedererkennen durch Ohrenzeugen im Rahmen der Gegenüberstellung aus der Perspektive der Ermittlungsbeamten (und des Sachverständigen) als (versuchtes) „Identifizieren“ im Sinne einer objektbezogenen Analyse darstellt: Vergleichsobjekte sind hier die (persönliche Wahrnehmung der) „Stimme einer tatverdächtigen Person“ und die „Aussage eines Ohrenzeugen zur Täterstimme“. Von daher ist das Bemühen der Vernehmungsbeamten verständlich (und auch erforderlich), Kriterien für die eigene (analytische) Beurteilung zu erfragen, weil diese Kriterien für die Ermittlungsbeamten die einzige Möglichkeit sind, sich selbst „ein Bild zu machen“ und die Angaben des Ohrenzeugen zu kontrollieren. Kann die Beweisperson jedoch (mangels sprachlicher Gewandtheit oder akustischen Differenzierungsvermögens oder aus anderen Gründen) keine Kriterien nennen, erschwert dies zwar die Kontrollmöglichkeit der Ermittlungsbeamten (und des Sachverständigen), dürfte aber für sich genommen kein Grund sein, die Aussage der Beweisperson pauschal als unzuverlässig zu verwerfen, da sich das Wiedererkennen als ganzheitlicher Wahrnehmungsprozeß in aller Regel einer verstandesmäßigen (intellektuellen) Begründung entzieht.

Stellt sich demnach gegenüber dem „Wiedererkennen“ seitens der Beweisperson das Bemühen der Vernehmungsperson nach Kontrolle der Identitätsaussage als „Identifizieren“ dar, werden hinsichtlich des Identitätsnachweises ganz grundlegende Unterschiede in der „Erkennungsstrategie“ von Vernehmungsperson und Beweisperson erkennbar. Wer diese Unterschiede als Vernehmungsperson (oder als Sachverständiger) nicht klar erfaßt und differenziert in seine praktische Arbeit einfließen läßt, wird die Angaben eines Ohrenzeugen nur unzureichend würdigen können.

3/5 Fehlerquellen

Im folgenden werden einige Fehlerquellen des Wiedererkennens von Stimmen behandelt (siehe dazu allgemein „Identifizieren und Wiedererkennen“, „Gegenüberstellung zwecks Wiedererkennens“, „Lichtbildvorlage“, „Wiedererkennen, wiederholtes“, „Zeugenaussage“, „Vernehmung“ und „Fragetechnik“).

Als Faustregel wird folgende Annahme formuliert: Je charakteristischer die Täterstimme, je länger die Konfrontation einer Person mit dieser Stimme in der Tatsituation, je besser die akustische Qualität des zu beurteilenden Vergleichsmaterials und je größer die Anzahl der Beurteilungspersonen, desto geringer wird die Fehlerwahrscheinlichkeit beim Prozeß des Wiedererkennens sein [Künzel 1989, 838].

3/5.1 Wahrnehmung

Die akustische Wahrnehmung in der Tatsituation hängt ganz grundlegend von der Hörfähigkeit der Beweisperson ab; das wird vor allem bei älteren Personen nicht selten übersehen: Die Wahrnehmung hoher Frequenzen verschlechtert sich mit zunehmenden Alter [Bender 195]. Durch entsprechendes Befragen und Beobachten kann bis zu einem gewissen Grade auch während der Hauptverhandlung ermittelt werden, ob eine Person schwerhörig ist (überdurchschnittlich lautes Sprechen, häufiges Nachfragen bei gestellten Fragen, Fehler bei der Aussprache bestimmter Laute) oder ob sie ein Hörgerät trägt [GA 1988, 220]. Das Hörvermögen eines Ohrenzeugen sollte selbst dann von einem Audiologen untersucht werden, wenn die Beweisperson scheinbar nur eine geringe oder gar keine Hörschädigung hat, vor allem aber dann, wenn die Aussage von entscheidender Bedeutung ist [Hammersley 128].

Ein zuverlässiges Wiedererkennen setzt voraus, daß die Beweisperson die wiederzuerkennende Stimme in der Tatsituation mindestens eine Minute gehört hat; nach anderen Studien sollen allerdings auch wenige Sekunden ausreichen [Bender 193].

Neben der Wahrnehmungsdauer ist der Inhalt der Kommunikation („Vielfalt der gehörten linguistischen Reize“) von Bedeutung; auch insoweit sind Aufmerksamkeit und Interesse der

Beweisperson wesentliche Bedingungen ihrer Wiedererkennungseistung [Bender 194]. Das Gelingen eines späteren Wiedererkennens von Stimmen hängt damit entscheidend von den Umständen ab, unter denen sie gehört wurden [Eisenberg 477].

Die Wahrnehmung von Stimmen in der Tatsituation kann von Streß (Aufregung, Angst) beeinflusst sein [Künzel 838]. Eine Klärung dieser Einflüsse im einzelnen gilt als zweifelhaft, da sich Menschen in vergleichbaren Extremsituationen durchaus unterschiedlich und gegensätzlich verhalten (können) [GA 1988, 222].

Hintergrundlärm in der Wahrnehmungssituation verschlechtert das Wiedererkennen von Stimmen, unter anderem wird auch die Hörfähigkeit der Beweisperson beeinträchtigt [Bender 194]. Das gilt insbesondere für Telefonübertragungen [GA 1988, 220]. In der Tatsituation wird der Täter deshalb möglicherweise lauter sprechen, so daß sich der Klang seiner Stimme verändert [Bender 194]. Auch die Sichtbarkeit der sprechenden Person [Bender 194] sowie ihre Entfernung [Eisenberg 477] können einen Einfluß auf die Wahrnehmung von Stimmen haben: Die visuelle Wahrnehmung des Sprechers (in der Tatsituation) kann die akustische Wahrnehmung beeinflussen (und umgekehrt) [Bender 194].

Ein Verstellen der Stimme in der Wahrnehmungssituation kann die Wiedererkennungseistung erheblich verschlechtern; selbst einfache Veränderungen haben starke Auswirkungen [Bender 194].

Beispiel: Langsames Sprechen, Flüstern [Bender 194].

Das gilt auch für mehr oder weniger „natürliche“ Veränderungen der Stimme.

Beispiel: Heiserkeit, Streß, Affekt [Bender 194].

Gefühle und Streß verändern unter anderem Tonhöhe und Lautstärke einer Stimme; vermutlich achten Hörer normalerweise auf Merkmale der Stimme, die Gefühle oder andere Bedeutungsinhalte vermitteln [Hammersley 127].

Bei der Wahrnehmung mehrerer unbekannter Stimmen kann es zu „Interferenzen“ (Überlagerung und gegenseitige Verstärkung oder Abschwächung) kommen [Bender 194]. Der „optimale

Ohrenzeuge“ sollte deshalb während der Tatausführung nur eine, nicht mehrere Stimmen gehört und besonders auf diese geachtet haben [Hammersley 133].

3/5.2 Erinnerung

Als einzige in sämtlichen wissenschaftlichen Untersuchungen bestätigte Tatsache wird der Umstand genannt, daß die Fähigkeit, andere Personen aufgrund ihrer Stimme zu erkennen, „individuell stark unterschiedlich“ ausgeprägt ist [GA 1988, 220].

Blinde und musikalische Personen vermögen Stimmen besser wiederzuerkennen als andere Personen; Kinder bis zu wenigstens zehn Jahren können Stimmen nicht so gut erkennen wie Erwachsene [Bender 195].

Zwischen dem Gedächtnis für Gefühle und demjenigen von Stimmen besteht möglicherweise eine Wechselwirkung [Hammersley 127]. Das Gedächtnis für vokale Informationen ist deutlich besser, wenn die Beweisperson ihre Aufmerksamkeit auf die sprachliche Information gerichtet hat als wenn eine Kommunikation nur beiläufig erfolgte [Bender 194].

Entgegen früheren Annahmen soll die Wiedererkennensleistung bezüglich menschlicher Stimmen nach neueren (nicht näher belegten) Studien zumindest einen Monat nach dem Erstkontakt noch keine Verminderung aufweisen „sondern im Gegenteil die Erkennungsrate leicht höher .. als unmittelbar nach dem Erstkontakt“ sein; der Gegensatz zwischen diesen Befunden wird „mindestens zum Teil auf wesentliche Unterschiede im Testablauf“ zurückgeführt [Künzel 837]. Das zeige im übrigen „wie genau eine jeweils im konkreten forensischen Einzelfall vorliegende Konstellation von Beurteilungsbedingungen zu analysieren ist, bevor eine Bewertung der Erkennungsurteile erfolgen kann“ [Künzel 837]. „Ungeachtet der noch ungelösten Probleme ist es .. wahrscheinlich, daß größere Zeitintervalle die Wiedererkennensfähigkeit verschlechtern [Hammersley 124].

Bezüglich des Bekanntschaftsgrades weisen Testergebnisse darauf hin, daß das Wiedererkennen von Stimmen verwandter oder bekannter Personen stark von der Art des Sprachmaterials abhängt, und zwar insbesondere vom Vorhandensein gewohnter

persönlicher Begrüßungsfloskeln, dem angesprochenen Thema oder anderen, von der Sprechsituation abhängigen „pragmatischen Faktoren“ [Künzel 837]. Sind diese Voraussetzungen gegeben, besitzt die Mehrheit der Testhörer einen Vertrauensvorsprung gegenüber mit den Sprechern nicht bekannten Testhörer; besteht die Beurteilungsgrundlage demgegenüber aus einem neutralen Text, wie etwa einer Menge von Testsätzen oder einem kurzen Zeitungsausschnitt, ergeben sich keine statistisch signifikanten Unterschiede zwischen den Wiedererkennensleistungen beider Klassen von Testhörern [Künzel 838]. Dessen ungeachtet kann das Beurteilungsverhalten des einzelnen Testhörers in beiden Klassen vom Verhalten der Mehrheit zur einen oder anderen Seite „stark“ abweichen [Künzel 838].

In den Fällen akustischer Gegenüberstellung ergibt sich der Bekanntschaftsgrad von Stimmen in der Regel ausschließlich aus der Tatsituation und ist daher im allgemeinen sehr gering; ist die Täterstimme dem Ohrenzeugen schon von früher bekannt, kommt ein Wiedererkennen schon allein aufgrund oder gar in der Tatsituation in Betracht, so daß eine akustische Präsentation von Vergleichsmaterial zum Zwecke des Identitätsnachweises regelmäßig deshalb ausscheidet, weil der Ohrenzeuge sich bereits zur Identität der fraglichen Person geäußert hat.

3/5.3 Wiedergabe

Gibt die Beweisperson für ihr Wiedererkennen besondere Merkmale an, ist stets damit zu rechnen, daß sie auch andere Personen mit diesen Merkmalen „wiedererkennt“.

Wird als Merkmal ein bestimmter Dialekt genannt, ist zu berücksichtigen, daß das tatsächliche Vorhandensein eines Dialektes in aller Regel von dem subjektiven Eindruck des Hörers abhängt [Hammersley 125].

Beispiel: Während für einen Norddeutschen alle bayerischen Dialekte gleich klingen, kann ein Bayer durchaus in der Lage sein, Stimmen mit bayerischem Dialekt aufgrund regionaler Besonderheiten zu unterscheiden.

Ein fremder Dialekt kann dazu beitragen, die Täterstimme von den Stimmen anderer Personen mit vertrautem Dialekt zu

unterscheiden; es besteht jedoch die Gefahr, daß der Hörer später jeden Sprecher mit einem ähnlichen Dialekt fälschlicherweise als Täter „identifiziert“ [Hammersley 125].

Für Stimmen von Ausländern ist die Wiedererkennensleistung im allgemeinen schlechter als für Stimmen von einheimischen Personen [Hammersley 125].

Unterschiedliche Sprachen und Jargons werden auch in verschiedenen Generationen derselben Sprachgemeinschaft gesprochen, so daß sich für ältere Leute alle jungen Stimmen gleich anhören und umgekehrt [Hammersley 128].

Trägt die Erinnerung an eine ungewöhnliche Stimme dazu bei, daß die Beweisperson eine bestimmte Person als Träger dieser Stimme „identifiziert“, ist immer mit der Möglichkeit zu rechnen, daß es irgendeine andere Person mit ähnlicher Stimme ist, selbst wenn sich die Beweisperson subjektiv ganz sicher ist; die „identifizierte“ Person mag am Ende „der einzige Berliner, Amerikaner oder Stotterer sein, den der Zuhörer kennt“ [Hammersley 125].

Soweit der Akzent oder die Einzigartigkeit einer Stimme für die Beweiswürdigung von Bedeutung ist, sollte der Sachverständige eines Linguisten oder phonetischen Sachverständigen konsultiert werden [Hammersley 125].

3/5.4 Gegenüberstellung

In den forensisch bedeutsamen Fällen sind Gegenüberstellungen zum Zwecke des Wiedererkennens von Stimmen in der Regel Bestandteil zeugenschaftlicher Vernehmungen. Als Fehlerquellen sind daher grundsätzlich sämtliche Umstände in Betracht zu ziehen, die auch das Ergebnis von Vernehmungen, visuellen Gegenüberstellungen, Lichtbildvorlagen oder das wiederholte Wiedererkennen beeinflussen (siehe dazu im einzelnen die entsprechenden Stichworte).

Bei der Vorbereitung und Durchführung akustischer Gegenüberstellungen sind die für visuelle Gegenüberstellungen entwickelten Grundsätze zu beachten [Bender 195].

3/5.4.1 Voraussetzungen

Eine akustische Gegenüberstellung kommt grundsätzlich nur dann in Betracht, wenn die Umstände der Tatsituation, insbesondere die akustischen Bedingungen (Telefon-, Funkübertragung oder direkte persönliche Konfrontation) und die Dauer des sprachlichen Kontakts (Expositionsdauer) die Durchführung eines Tests sinnvoll erscheinen lassen („Randbedingungen“): Weist die Täterstimme nach Angaben der Beweisperson keine unmittelbar erkennbaren Besonderheiten (wie etwa Heiserkeit oder Lispeln) auf, ist eine Exposition von einer Minute als Minimum anzusehen; liegt bereits nach dem Eindruck der Beweisperson eine offensichtliche Verstellung der Täterstimme vor, kommt eine Gegenüberstellung überhaupt nicht in Betracht, da angesichts der vielfältigen Möglichkeiten der Verstellung von sprachwissenschaftlich ungeschulten Personen nur wenige Stimmverstellungen aufgedeckt und angemessen beschrieben werden können [GA 1988, 224].

Mit der Durchführung einer akustischen Gegenüberstellung sollte nicht unnötig gewartet werden; Zeitverzögerung kann die Wiedererkennensleistung beeinträchtigen [Hammersley 130]. Zudem wirken sich unbewußte Irreführungen aller Art im allgemeinen stärker aus, wenn die Erinnerung verblaßt und die Beweisperson unsicher wird [Hammersley 130].

Das Wiedererkennen von Stimmen sollte im Rahmen einer rein akustischen Gegenüberstellung erfolgen: die Beweisperson sollte die zu den Stimmen gehörenden Personen nicht sehen können [Bender 195], da gleichzeitiges Sehen die Wiedererkennensleistung beeinflusst [Bender 194]. Die Präsentation des Vergleichsmaterials mittels Tonbandaufnahmen ist der akustischen „Realgegenüberstellung“ vorzuziehen [Bender 195].

Die technischen Bedingungen für die Durchführung der Gegenüberstellung müssen optimal sein [GA 1988, 224]. Die Tonbandaufnahmen sollten von bester technischer Qualität sein (HiFi-Kassettenrecorder mit hochwertigem Mikrophon statt Diktaphon oder Kassettenrecorder mit eingebautem Mikrophon); Aufnahmen schlechter Qualität können die Wiedererkennensleistung vermindern [Hammersley 130].

Die Beweisperson sollte zumindest gefragt werden, ob sie die Stimmen auf dem Tonband hören kann [Hammersley 130];

bestehen Zweifel an der Hörfähigkeit der Beweisperson, ist vorab ein Hörtest durchzuführen [Bender 195].

3/5.4.2 Vergleichsstimmen

Der Beweisperson sollten mindestens zwanzig Stimmen präsentiert werden [Bender 195]; bei einer geringeren Anzahl von Stimmen ist das Risiko, daß der Verdächtige fälschlicherweise durch bloßes Raten „wiedererkannt“ wird, vergleichsweise hoch [Hammersley 131].

Beispiel: Werden nur fünf Stimmen vorgegeben, wird die Wahl mit einer Wahrscheinlichkeit von 20% rein zufällig auf einen Sprecher fallen; die höchste „in der Wissenschaft akzeptable Irrtumswahrscheinlichkeit“ liegt bei 5% (1/20) [Hammersley 131].

Alle Vergleichsstimmen sollten Ähnlichkeiten in wesentlichen Merkmalen aufweisen [Bender 195].

Beispiel: Geschlecht, Alter, Volumen, Dialekt, Hintergrundgeräusche [Bender 195].

Der Tatverdächtige sollte davon abgehalten werden, seine Stimme während der Tonbandaufnahmen zu verstellen [Hammersley 131].

Um sicherzustellen, daß sich die Stimme eines Tatverdächtigen nicht rein akustisch von den übrigen Vergleichsstimmen abhebt, sollte das Material vorab anderen Personen vorgespielt und entsprechend getestet werden [Bender 195]. Eine besondere Gefahr akustischer Gegenüberstellungen wird nämlich darin gesehen, daß die Beweisperson im Vorfeld der Gegenüberstellung bestimmte sprachliche Merkmale weder erkannt noch beschrieben oder erwähnt hat und bei nachlässiger Abstimmung der übrigen Vergleichsstimmen auf die Stimme des Tatverdächtigen zu besorgen ist, daß letztere allein durch ihre eigentümlichen Merkmale hervorsticht [Eisenberg 478].

Beispiel: Für einen bayerischen Polizeibeamten, der eine Gegenüberstellungsgruppe zusammenstellt, klingen möglicherweise alle norddeutschen Dialekte gleich, während sein Kollege aus Hamburg noch eine Stimme heraushören kann, die sich durch

ihren spezifischen Akzent von den übrigen abhebt; umgekehrt können sich alle Stimmen mit bayerischem Dialekt für einen Norddeutschen gleich anhören, während ein Hörer aus der jeweiligen Region durchaus einzelne Sprecher mit bayerischem Dialekt zu unterscheiden vermag [Hammersley 125].

Die „fatale Fehlermöglichkeit“, daß eine als „Täterstimme“ bezeichnete Stimme eines Tatverdächtigen nicht mit der tatsächlichen Täterstimme identisch ist, gleichwohl aber bei der akustischen Gegenüberstellung konsistent und ausschließlich für diese gehalten wird, weil sie gemäß der Beschreibung der Beweisperson aufgrund einer zufällig vorhandenen mehr oder weniger großen Ähnlichkeit von den Ermittlungsbeamten als mögliche Täterstimme in Betracht gezogen wurde, kann nur durch eine „sorgfältige Vorauswahl“ der zur akustischen Gegenüberstellung verwendeten Stimmen vermieden werden [GA 1988, 219].

Das Vorhandensein ungewöhnlicher stimmlicher Merkmale erhöht zwar die Wahrscheinlichkeit eines „richtigen“ Wiedererkennens, soweit beim Hörtest auch die Täterstimme präsentiert wird; wird statt der Täterstimme jedoch eine andere Stimme präsentiert, die die markanten Merkmale in einer ähnlichen Ausprägung (zufällig) ebenfalls aufweist, erhöht sich die Gefahr eines „falschen“ Wiedererkennens [Eisenberg 478]. Das „nicht mit letzter Sicherheit“ auszuräumende Risiko „faktische(r) Nichtidentität“ kann auch experimentell nicht aufgedeckt werden [GA 1988, 219].

3/5.4.3 Vergleichstext

Die zu Vergleichszwecken gesprochenen Texte sollten die Formulierungen enthalten, die der Täter tatsächlich gesprochen hat; dabei sollten auch bekannte Eigenheiten der Betonung und Aussprache reproduziert werden [Bender 195]. Ein veränderter sprachlicher Kontext erschwert das Wiedererkennen [Hammersley 131].

Mehr als ein kurzer Satz von jeder Stimme scheint nicht erforderlich zu sein: Erkennt die Beweisperson nämlich überhaupt die Stimme wieder, genügt dafür bereits ein Satz von etwa acht bis dreißig Worten [Hammersley 130].

3/5.4.4 Personal

Das mit der Durchführung der akustischen Gegenüberstellung betraute Personal sollte selbst nicht wissen, welches die Stimme des Tatverdächtigen ist [Bender 195]. Anderenfalls ist nicht auszuschließen, daß sie die Beweisperson unbewußt auf diese Stimme hinweisen; solche Hinweise können zum fälschlichen „Wiedererkennen“ des Verdächtigen führen [Hammersley 132].

3/5.4.5 Ablauf

Vor der akustischen Gegenüberstellung sollten Ohrenzeugen nicht in Kontakt mit dem Tatverdächtigen kommen; anderenfalls besteht die Gefahr, daß die Stimme des Verdächtigen aufgrund dieses Kontaktes wiedererkannt wird und nicht deshalb, weil der Verdächtige tatsächlich der Täter ist [Hammersley 130]. Schon der alleinige Anblick des Verdächtigen kann das anschließende Wiedererkennen beeinflussen, da im allgemeinen von der Stimme auf die äußere Erscheinung der sprechenden Person geschlossen werden kann [Hammersley 130]. Ohrenzeugen sollten sich vor der Gegenüberstellung auch nicht mit anderen Beweispersonen oder Bekannten des Verdächtigen besprechen [Hammersley 130].

Die Beweisperson ist darauf hinzuweisen, daß die Stimme des Täters sich möglicherweise nicht unter den präsentierten Vergleichsproben befindet [Bender 195] und daß die Beweisperson nicht etwa irgendeine Stimme benennen sollte, wenn sie die Stimme des Täters nicht hört; die Beweispersonen sollten sich nicht gedrängt fühlen, eine Stimme auszuwählen [Hammersley 132].

Vor einer Entscheidung sollten der Beweisperson zunächst sämtliche Stimmen vorgespielt werden; auf Wunsch sollte dies beliebig oft wiederholt werden [Bender 195]. Die Beweisperson sollte darauf hingewiesen werden, daß sie sich das Tonband beliebig oft anhören kann, ehe sie eine endgültige Entscheidung trifft; diese Entscheidung sei jedoch nur einmal möglich [Hammersley 132]. Es kann hilfreich sein, die Beweisperson anhand einer Liste die ihr ähnlich erscheinenden Stimmen auswählen zu lassen und ihr weiterhin nur noch diese Stimmen vorzuspielen [Hammersley 132].

Hat sich die Beweisperson statt für den Tatverdächtigen für eine andere Vergleichsperson entschieden oder kann sie sich überhaupt nicht entscheiden, sollte die Gegenüberstellung abgebrochen werden [Bender 195]; mit einer Fortsetzung des Wiedererkennens wird die Gefahr falscher Ergebnisse immer wahrscheinlicher, weil sich die Beweisperson entweder nicht genau an die Stimme des Täters erinnern kann oder der Verdächtige tatsächlich nicht der Täter ist [Hammersley 132].

3/5.4.6 Wiederholung

Hat sich die Beweisperson für eine Stimme entschieden, kann sie bei einem zweiten Durchgang oder in einer erneuten Gegenüberstellung diese Stimme nur deshalb „wiedererkennen“, weil sie sich an die erste Gegenüberstellung erinnert; einen stärkeren Beweiswert des Wiedererkennens bedeutet dies nicht [Hammersley 132].

3/5.4.7 Dokumentation

Sämtliche Einzelheiten der Vorbereitung und Durchführung der akustischen Gegenüberstellung sind aktenkundig zu machen; die Akten müssen insbesondere die verwendeten Tonaufzeichnungen enthalten [Bender 195].

Ganz entscheidend ist die exakte Protokollierung der Angaben, welche die Beweisperson vor und nach der Gegenüberstellung macht, gegebenenfalls der Verhaltensregeln und Einweisung der Beweisperson sowie des genauen Vorgehens der Polizeibeamten im Rahmen etwaiger „Sofortmaßnahmen“ [Kriminalistik 1995,656].

3/6

Beweiswert

Der Beweiswert des Wiedererkennens von Stimmen sollte mit Zurückhaltung betrachtet werden: Während unter bestimmten Bedingungen die Aussage des Ohrenzeugen „durchaus sehr zuverlässig sein“ kann, können „geringfügig andere Umstände“ die Aussage völlig wertlos machen [Hammersley 114]; selbst unter optimalen Bedingungen sind Hörer aber nicht immer in der Lage, unbekannte Personen allein anhand ihrer Stimme zu erkennen [Hammersley 133].

Das Wiedererkennen von Stimmen kann auch von nicht-akustischen Parametern, insbesondere der Art und Weise, wie die Stimme ursprünglich wahrgenommen wurde und wie sie bei der akustischen Gegenüberstellung dargeboten wird, beeinflusst werden [Hammersley 133]. Sind die Variationen in der Stimme eines Sprechers größer als die Unterschiede zwischen den Stimmen mehrerer Sprecher, kann ein Wiedererkennen unmöglich sein, was vor allem in Frage kommt, wenn die Beweisperson den Sprecher und das Spektrum seines sprachlichen Ausdrucks nicht kennt und „keine Hinweise aus dem Kontext“ heranziehen kann [Hammersley 133].

Ohrenzeugen können sich vielleicht an bestimmte Aspekte einer Stimme erinnern, doch reicht dies möglicherweise nicht aus, um eine Person zuverlässig an ihrer Stimme wiederzuerkennen [Hammersley 133].

Zur Überprüfung des Beweiswertes akustischer Gegenüberstellungen kommt die Beauftragung eines Sachverständigen in Betracht [GA 1988, 218].

Beispiel: Anhörung eines Sachverständigen zur Frage, „... ob die .. aufgezeichneten Stimmen .. nach Inhalt, Umfang und technischer Ausgestaltung geeignet sind, .. einer Zeugin für einen zwei Monate zurückliegenden Vorfall als verlässlicher Vergleich zu einer – seinerzeit nicht auf Tonträger festgehaltenen – Stimme zu dienen und welche Fehlermöglichkeiten bestehen“ [GA 1988, 218].

Dabei ist allerdings zu beachten, daß jeder Versuch, die Wahrnehmungsleistung des Ohrenzeugen in wissenschaftlich-analytischen Kategorien objektbezogener Analysen zu beurteilen, den

ganzheitlichen Charakter akustischer Wahrnehmung und Erinnerung vernachlässigt. Soweit daher sachverständigerseits das Wiedererkennen durch Ohrenzeugen als „Sprechererkennung durch Laien“ [Künzel 836] oder „Heranziehung des wissenschaftlichen Laien als Stimmenerkener“ [GA 1988, 215] problematisiert oder der Frage nachgegangen wird, „ob Sprecheridentifizierung durch sprachwissenschaftliche Laien für forensische Zwecke taugt oder nicht“ [GA 1988, 223], kann durchaus zweifelhaft sein, ob aufgrund der in diesen Formulierungen sichtbar werdenden wissenschaftlich-analytischen Herangehensweise die Probleme ganzheitlicher akustischer Wahrnehmung von Ohrenzeugen angemessen und zuverlässig erfaßt werden (Sind Sprachanalytiker als solche unter forensischen Bedingungen grundsätzlich die zuverlässigeren Wiedererkennenszeugen?); von derartigen fachlichen Grundanschauungen kann aber gerade auch der Beweiswert sachverständiger Ausführungen maßgeblich beeinflusst sein (siehe dazu „Sachverständigengutachten“).

3/7 Rechtsprechung

Zur abweichenden Terminologie (Wiedererkennen als Fall des Identifizierens) siehe „Identifizieren und Wiedererkennen“.

3/7.1 Wahlgegenüberstellung

Für die Identifizierung eines Tatverdächtigen aufgrund eines Stimmvergleichs gelten die für die Gegenüberstellung mit einem Augenzeugen anerkannten Grundsätze entsprechend¹.

Es ist sicherzustellen, daß der Zeuge die Stimme des Verdächtigen nicht isoliert sondern neben anderen Stimmen hört². Die Vergleichsstimmen müssen eine gewisse Klangähnlichkeit aufweisen; es dürfen dem Zeugen neben dem mit einem fremdländischen Akzent oder einem Dialekt sprechenden Verdächtigen nicht etwa Stimmen einer anderen Sprachheimat vorgestellt werden³.

Daß ein Zeuge neben der Stimme des Angeklagten auch die eines Kriminalbeamten hören kann, eröffnet keine wirkliche Wahlmöglichkeit⁴. Das gilt auch für den Fall, daß die Zeugen lediglich ein Gespräch zwischen dem Angeklagten und zwei Kriminalbeamten verfolgen können, zumal dann, wenn nur der Angeklagte die bei der Tat von den Zeugen gehörten Sätze spricht⁵.

3/7.2 Dokumentation

Ist der bei der Polizei durchgeführte Stimmvergleich nicht auf Tonträger festgehalten worden, hat sich das Tatgericht (gegebenfalls

¹ BGH Urteil v. 24.02.1994 – 4 StR 317/93 = BGHSt 40, 66 = NSiZ 1994, 295 = StV 1994, 282 m. Anm. Achenbach/Perschke 577; BGH Beschluß v. 12.04.1994 – 4 StR 142/94 = NSiZ 1994, 597 m. Anm. Eisenberg = StV 1994, 638 (Wiedererkennen – R 2 –).

² BGH Urteil v. 24.02.1994 – 4 StR 317/93 = BGHSt 40, 66 = NSiZ 1994, 295 = StV 1994, 282 m. Anm. Achenbach/Perschke 577; BGH Beschluß v. 12.04.1994 – 4 StR 142/94 = NSiZ 1994, 597 m. Anm. Eisenberg = StV 1994, 638 (Wiedererkennen – R 2 –).

³ BGH Urteil v. 24.02.1994 – 4 StR 317/93 = BGHSt 40, 66 = NSiZ 1994, 295 = StV 1994, 282 m. Anm. Achenbach/Perschke 577.

⁴ BGH Urteil v. 24.02.1994 – 4 StR 317/93 = BGHSt 40, 66 = NSiZ 1994, 295 = StV 1994, 282 m. Anm. Achenbach/Perschke 577.

⁵ BGH Beschluß v. 12.04.1994 – 4 StR 142/94 = NSiZ 1994, 597 m. Anm. Eisenberg = StV 1994, 638 (Wiedererkennen – R 2 –).

nenfalls unter Hinzuziehung eines Sprachwissenschaftlers) durch Anhören der Vergleichsstimmen selbst einen Eindruck hinsichtlich des Beweiswertes des Stimmvergleichs zu verschaffen¹.

3/7.3 Beweiswert

Einer Identifizierung durch Ohrenzeugen kann nicht etwa deshalb der Beweiswert abgesprochen werden, weil die als Täter in Betracht kommende Person keinen Sprachfehler hat; auch geringer ausgeprägte Sprachmerkmale können zur Identifizierung ausreichen². Welche Merkmale eine Identifizierung noch zuverlässig ermöglichen können, ist eine Frage des Einzelfalles, die der Tatrichter gegebenenfalls durch Hinzuziehung eines Sprachwissenschaftlers zu beantworten hat³.

Bei Mängeln des Stimmvergleichstests verliert die Identifizierung der Stimme durch den Zeugen nicht notwendig jeden Beweiswert; wie bei der fehlerhaften visuellen Gegenüberstellung muß sich der Tatrichter aber des besonderen Risikos einer Falschidentifizierung (erkennbar) bewußt sein⁴.

Ob die fehlerhafte Durchführung eines Stimmvergleichs den Beweiswert des Wiedererkennens der Stimme im Einzelfall möglicherweise so wenig berührt, daß eine fehlende Erörterung in den Urteilsgründen nicht zu beanstanden wäre, etwa weil der Zeuge bestimmte Sprachfehler oder Eigenarten schildert, die so auffällig sind, daß sie eine zuverlässige Identifizierung ermöglichen, bedarf dann keiner Entscheidung, wenn sich die vom Tatgericht nicht näher beschriebene „Stimmlage“ und (jedenfalls bei Taten, die in der Pfalz begangen werden) der „Pfälzer Dialekt“ („Gib mers Geld!“, „Das im Bodetresor ach!“, „Geld her, oder es knallt, aber alles!“) nicht als derart auffällige Eigenarten darstellen⁵.

¹ OLG Köln, Urteil v. 11.06.1996 – Ss 194/96 = NSiZ 1996, 509.

² OLG Köln, Urteil v. 11.06.1996 – Ss 194/96 = NSiZ 1996, 509.

³ OLG Köln, Urteil v. 11.06.1996 – Ss 194/96 = NSiZ 1996, 509 („kölscher Dialekt, undeutlich, schwammig, holprig oder ähnlich“).

⁴ BGH Urteil v. 24.02.1994 – 4 StR 317/93 = BGHSt 40, 66 = NSiZ 1994, 295 = StV 1994, 282 m. Anm. Achenbach/Perschke 577; BGH Beschluß v. 12.04.1994 – 4 StR 142/94 = NSiZ 1994, 597 m. Anm. Eisenberg = StV 1994, 638 (Wiedererkennen – R 2 –).

⁵ BGH Beschluß v. 12.04.1994 – 4 StR 142/94 = NSiZ 1994, 597 m. Anm. Eisenberg = StV 1994, 638 (Wiedererkennen – R 2 –).

Hat ein Zeuge die Stimme des Angeklagten schon bei verschiedenen früheren Begegnungen gehört, muß daraus nicht gefolgert werden, daß er diese Stimmen schon bei der Tat hätte erkennen müssen; denkbar ist aber, daß sich bei einer „Gegenüberstellung“, die infolge der Heimlichkeit des Mithörens eines Gesprächs zwischen dem Angeklagten und einem Kriminalbeamten in hohem Maße suggestiv wirkt, die (blasse) Erinnerung des Zeugen an die Stimme des Angeklagten und die an die Stimme des Täters vermengen und er nicht, wie er meint, diese, sondern jene wiedererkennt¹.

3/7.4 Beweis Antrag

Läuft ein Beweis Antrag auf Durchführung eines Stimmvergleichstests letztlich auf eine Wiederholung einer bereits durchgeführten Beweiserhebung unter abweichenden (nicht verlässlichen) Bedingungen hinaus, darf das Tatgericht davon ausgehen, daß es dem Ergebnis einer solchen Beweiserhebung nicht folgen werde, und den Beweis Antrag als „für die Entscheidung ohne Bedeutung“ ablehnen². Das gilt insbesondere für den Fall, daß ein Sachverständiger einen unter nicht optimalen Bedingungen tatnah und „live“ durchgeführten Stimmvergleichstest als „durchaus beachtlich“ und zur Wiedererkennung geeignet bewertet und die Auffassung vertritt, daß der beantragte neue Test wegen ungünstiger Voraussetzungen (leichte Verfälschung der Stimmen durch Bandaufnahme, Sprechen nur kurzer einzelner Sätze und Tatferne) nicht zu besseren Beweisergebnissen führen würde³.

¹ BGH Urteil v. 24.02.1994 – 4 StR 317/93 = BGHSt 40,66 = NStZ 1994,295 = StV 1994,282 m. Anm. Achenbach/Perschke 577.

² BGH Beschluß v. 03.09.1996 – 1 StR 426/96 = NStZ 1997, 95.

³ BGH Beschluß v. 03.09.1996 – 1 StR 426/96 = NStZ 1997, 95.

3/8 Resümee

Das Wiedererkennen wurde anschaulich beschrieben als „Ergebnis einer schöpferischen Verarbeitung lückenhafter Wahrnehmungen“, als „eine Art von Theorie“ über das, was „wirklich“ (tatsächlich) erlebt wurde [NStZ 82, 195] – womit neben dem inneren Vorgang des Wiedererkennens zugleich auch seine Problematik sowie die Schwierigkeiten einer nachvollziehenden Kontrolle anschaulich skizziert werden.

Angesichts der hohen Gefahr suggestiver Beeinflussung in der Vernehmungssituation unterliegen auch akustische Gegenüberstellungen den für die Durchführung und Beurteilung visueller Gegenüberstellungen entwickelten Grundsätzen; das gilt insbesondere für das Erfordernis einer ausreichenden Wahlmöglichkeit.

Da in der praktischen Polizeiarbeit nicht selten die exakte Beachtung der für Gegenüberstellungen erarbeiteten Standards vernachlässigt werden, stellen Rekonstruktion und Überprüfung einschlägiger Ermittlungen in aller Regel erfolversprechende Ansatzpunkte der Verteidigung dar.

4

Übersicht

Stimmvergleich		
	Identifizieren	Wiedererkennen
Erkenntnisquelle:	objektbezogene Analyse	subjektive Erinnerung
Erkenntnisvorgang:	Merkmalvergleich	Typisierendes Erleben
Hilfsmittel:	Tonaufzeichnungen	Tonaufzeichnungen Sprechproben Videoaufnahmen
Kompetenz:	Sachkunde	Bekanntschaft
Überprüfbarkeit:	intersubjektiv reproduzierbares Ergebnis	subjektive Behauptung
Fehlerquellen:	Material Methode Erfahrungswissen Qualifikation	Wahrnehmung Erinnerung Wiedergabe Gegenüberstellung