

Untersuchung zur Stimmerkennung bei aphasischen und nicht-aphasischen Patienten

Dissertation zur Erlangung des akademischen Grades eines
Doktors der Philosophie (Dr. phil.)

eingereicht von

Oliver Kneidl

Fakultät für Linguistik und Literaturwissenschaft
Universität Bielefeld

Gutachter:

HD Dr. Martina Hielscher-Fastabend
Professor Dr. Gert Rickheit

09. August 2006

Danksagung

Mein Dank gilt zunächst Frau HD Dr. Martina HIELSCHER-FASTABEND und Herrn Professor Dr. Gert RICKHEIT für die überaus wertvolle fachliche und menschliche Unterstützung zu allen Zeitpunkten meines Promotionsstudiums. Darüber hinaus danke ich Herrn apl. Professor Dr. med. Christoph J.G. LANG der Friedrich-Alexander-Universität Erlangen-Nürnberg (Neurologische Klinik mit Poliklinik) für die Überlassung der Initialidee dieser Arbeit. Dank auch allen Kolleginnen und Kollegen der Fakultät für Linguistik und Literaturwissenschaft (Abteilung Klinische Linguistik) der Universität Bielefeld, insbesondere Herrn Dr. Lorenz „Max“ SICHELSCHEMIDT für die Unterstützung bei der statistischen Auswertung, Frau Grainne DELANY für die vielen Dienste und Gefälligkeiten (auch außerhalb jeglicher Dienstzeiten und -vorschriften) sowie meiner Kollegin Frau Dr. Marion WITTLER für zahlreiche Diskussionen und fachliche Anregungen.

Ebenso danke ich der DEUTSCHEN FORSCHUNGSGEMEINSCHAFT (DFG), die durch die Bewilligung eines Doktorandenstipendiums im Graduiertenkolleg *Aufgabenorientierte Kommunikation* (GK 256) die Realisierung der Promotion finanziell ermöglichte.

Für hilfreichen fachlichen Austausch danke ich Frau Professorin Dr. Diana Roupas VAN LANCKER SIDTIS der New York University (Department of Speech-Language Pathology & Audiology) und Herrn Dr. med. Charly GAUL der Martin-Luther-Universität Halle-Wittenberg (Klinik und Poliklinik für Neurologie).

Besonderen Dank schulde ich allen beteiligten Probanden und Patienten sowie den Therapeuten der kooperierenden Kliniken in Bad Oeynhausen, insbesondere Frau Astrid SCHMIDKORT, Herrn Axel MÜLLER (beide Abteilung Sprachtherapie) und Herrn Dr. Dr. med. Erwin WEHKING (Chefarzt) der Klinik am Rosengarten sowie Herrn U. GEISLER (Abteilung für Neuropsychologie) und Herrn Dr. med. Michael LOEW (Chefarzt) der Klinik am Osterbach.

Nicht zuletzt möchte ich von ganzem Herzen allen kleinen und großen Menschen meiner näheren und weiteren Umgebung für die seelische Unterstützung während der vergangenen Jahre danken.

Oliver Kneidl
Bielefeld, im August 2006

Inhaltsverzeichnis

| | |
|-------------------------------------------------------------------------------------|-----------|
| 1. Einführung | 1 |
| 2. Begriffsbestimmung | 4 |
| 1. Theoretische Grundlagen | 7 |
| 3. Grundlagen der Phonetik und Akustik | 8 |
| 3.1. Artikulatorische Phonetik | 10 |
| 3.1.1. Atmung (Initiator) | 10 |
| 3.1.2. Phonation (Generator) | 11 |
| 3.1.3. Artikulation (Modifikator) | 14 |
| 3.2. Akustische Phonetik und psychoakustische Grundlagen | 24 |
| 3.2.1. Akustische Grundlagen der Schallproduktion | 24 |
| 3.2.2. Schallrezeption: Wahrnehmung von Tonhöhe und Lautstärke . | 28 |
| 4. Suprasegmentalia in der Spontansprache | 33 |
| 4.1. Erkennungsrelevante Sprechermerkmale | 35 |
| 4.2. Die Problematik der Inter- und Intra-variabilität von Stimmen | 36 |
| 4.3. Interindividuelle Merkmale der Stimme | 39 |
| 4.3.1. Die mittlere Sprechstimmlage | 39 |
| 4.3.2. Der Intonationsverlauf | 45 |
| 4.3.3. Die Stimmqualität | 46 |
| 4.4. Individuelle Merkmale der Sprache | 47 |
| 4.5. Individuelle Merkmale der Sprechweise | 47 |
| 5. Naive und professionelle Stimmerkennung | 49 |
| 5.1. Naive Stimmerkennung | 50 |
| 5.1.1. Einflüsse auf die Probandenleistung bei Sprechererkennungsaufgaben | 50 |

Inhaltsverzeichnis

| | |
|---------------------------------------------------------------------------|------------|
| 5.1.2. Sprechererkennung im Alltag | 62 |
| 5.2. Professionelle Sprechererkennung in der Forensik | 67 |
| 6. Anatomie und Physiologie der peripheren und zentralen Hörorgane | 71 |
| 6.1. Die peripheren Anteile des Ohres | 72 |
| 6.1.1. Das Außenohr (<i>Auris externa</i>) | 72 |
| 6.1.2. Das Mittelohr (<i>Auris media</i>) | 72 |
| 6.1.3. Das Innenohr (<i>Auris interna</i>) | 73 |
| 6.2. Primäre und sekundäre Hörrinde | 76 |
| 7. Neurolinguistische und -psychologische Grundlagen | 78 |
| 7.1. Grundlagen der Aphasie | 79 |
| 7.1.1. Standardsyndrome | 80 |
| 7.1.2. Sonderformen | 82 |
| 7.2. Störungen der Objekterkennung (<i>Agnosien</i>) | 84 |
| 7.3. Auditive Verarbeitungsstörungen | 85 |
| 7.3.1. Auditive Agnosien | 85 |
| 7.3.2. Phonagnosie | 88 |
| II. Experimenteller Teil | 105 |
| 8. Methodik | 106 |
| 8.1. Vorüberlegung zur Methodik | 107 |
| 8.2. Hypothesen | 108 |
| 8.3. Testdesign I: Vortestungen | 109 |
| 8.3.1. Anamnesegespräch | 110 |
| 8.3.2. Aachener Aphasie Test (AAT) | 111 |
| 8.3.3. Wechsler-Gedächtnistest – Revised (WMS-R) | 111 |
| 8.3.4. Alters-Konzentrations-Test (AKT) | 112 |
| 8.3.5. Hospital Anxiety and Depression Scale – Dt. Version (HADS-D) | 112 |
| 8.3.6. Durchführung I | 113 |
| 8.4. Testdesign II: Wahrnehmungsexperiment | 114 |
| 8.4.1. Verwendete Distraktorstimmen | 114 |
| 8.4.2. Beschreibung der Textstimuli | 115 |
| 8.4.3. Stichprobe | 117 |
| 8.4.4. Durchführung II | 117 |

Inhaltsverzeichnis

| | |
|------------------------------------------------------------------------------------------------------|------------|
| 9. Ergebnisse | 119 |
| 9.1. Stichprobe | 120 |
| 9.1.1. Gruppe 1: Aphasiker (LBD) | 120 |
| 9.1.2. Gruppe 2: Nicht-Aphasiker (RBD) | 120 |
| 9.1.3. Gruppe 3: gesunde Kontrollprobanden (HC) | 121 |
| 9.1.4. Zusammenfassung | 121 |
| 9.2. Zusammenfassung und Auswertung der Vortestergebnisse | 123 |
| 9.2.1. Ergebnisse der WMS-R und der HADS-D | 123 |
| 9.2.2. Übersicht der Prozentränge des AKT | 125 |
| 9.3. Statistische Überprüfung der Hypothesen | 126 |
| 9.3.1. Gesamterkennungsleistung | 127 |
| 9.3.2. Erkennungsleistungen auf Vokal-, CVC-, Wort- und Satzebene nach Stimulikategorie | 130 |
| 9.3.3. Erkennungsleistungen nach Stimulidauer | 134 |
| 9.4. Zusammenhänge zwischen Stimmerkennungsleistung und Vortest- ergebnissen | 136 |
| 9.4.1. Gruppe 1 (LBD) | 136 |
| 9.4.2. Gruppe 2 (RBD) | 136 |
| 9.4.3. Gruppe 3 (HC) | 137 |
| 9.4.4. Gesamtgruppe BD | 137 |
| 9.4.5. Gesamtgruppe | 137 |
| 9.5. Rating von Schwierigkeitsgrad und subjektiver Sicherheit | 138 |
| 9.5.1. Gruppe 1 (LBD) | 138 |
| 9.5.2. Gruppe 2 (RBD) | 140 |
| 9.5.3. Gruppe 3 (HC) | 142 |
| 9.5.4. Gesamtgruppe | 143 |
| | |
| 10. Zusammenfassung und Diskussion | 147 |
| | |
| 11. Ausblick | 160 |
| | |
| III. Literaturverzeichnis | 164 |
| | |
| IV. Anhang | 188 |
| | |
| 12. Tabellen und Abbildungen | 189 |
| 12.1. Tabellen | 190 |

Inhaltsverzeichnis

| | |
|----------------------------------------------------|------------|
| 12.2. Abbildungen | 199 |
| 13. Verwendete Test- und Untersuchungsbögen | 219 |
| 14. Verwendete Hard- und Software | 234 |
| 14.1. Verwendete Hardware | 235 |
| 14.2. Verwendete Software | 236 |

Abbildungsverzeichnis

| | |
|-------------------------------------------------------------------------------------------|-----|
| 3.1. Kehlkopf von ventral | 12 |
| 3.2. Kehlkopfknochen mit Glottis von oben ventral | 12 |
| 3.3. Sog- bzw. Bernoulli-Effekt | 13 |
| 3.4. Hohlräume im Ansatzrohr | 15 |
| 3.5. Stehende Schalldruckwellen im neutralen Rohr | 16 |
| 3.6. Sekundäre Kardinalvokale | 19 |
| 3.7. Vokaltrapez | 20 |
| 3.8. Pulmonale Konsonanten | 23 |
| 3.9. Spitzen- und effektiver Schalldruck | 26 |
| 3.10. Amplitudenveränderung über die Zeit | 27 |
| 3.11. Fourier-Synthese harmonischer Sinusschwingungen | 28 |
| 3.12. Geräusch als Zeitsignal und als kontinuierliches Spektrum | 28 |
| 3.13. Kurven gleicher Lautstärkepegel nach ISO 226 / ISO 226:2003 | 29 |
| 3.14. Zusammenhang zwischen Lautstärkepegel und Lautheit nach ISO 532B | 30 |
| 3.15. Zusammenhang zwischen subjektiver Tonhöhe und Frequenz | 31 |
| 3.16. Zusammenhang zwischen Bark und Frequenz | 31 |
| 4.1. Intraindividuelle Unterschiede am Beispiel <Ball> | 37 |
| 4.2. Zusammenhang zwischen Quellsignal und Filtercharakteristik | 42 |
| 4.3. Beispiel für einen Intonationsverlauf | 46 |
| 5.1. <i>Speaker recognition by listening</i> : Schema mit Sprecherinformationen | 50 |
| 6.1. Ohr (halbschematisch), teilweise eröffnet von frontal (rechts) | 73 |
| 6.2. Gehörknöchelchen, von oben medial (rechts) | 74 |
| 6.3. Knöchernes Labyrinth mit Cochlea) | 75 |
| 6.4. Heschl'sche Querwindung (<i>Gyrus temporalis transversus</i>) | 76 |
| 7.1. Läsionsorte bei aphasischen Syndromen | 84 |
| 7.2. Arbeitsmodell zur Personenerkennung | 101 |

Abbildungsverzeichnis

| | |
|-----------------------------------------------------------------------------|-----|
| 7.3. Arbeitsmodell zur Stimm- und Gesichtererkennung | 102 |
| 8.1. Methodenschema der Stimmerkennungsverfahren | 107 |
| 9.1. Alterszusammensetzung der Gruppen (n.s.) | 122 |
| 9.2. <i>p.-o.-time</i> der Patienten mit zerebraler Läsion (n.s.) | 122 |
| 9.3. Gruppenergebnisse in der WMS-R (n.s.) | 124 |
| 9.4. Gruppenergebnisse in der HADS-D (n.s.) | 124 |
| 9.5. Prozentrangmittelwerte R des AKT | 126 |
| 9.6. Gesamterkennungsleistung der Gruppen HC und BD | 127 |
| 9.7. Gruppenleistungen Gesamt (<i>sE + vE</i>) | 128 |
| 9.8. Leistungen auf V-, CVC-, W- und S-Ebene (<i>sE + vE</i>) | 131 |
| 9.9. Subjektiver Schwierigkeitsgrad (LBD) | 138 |
| 9.10. Subjektives Sicherheitsgefühl (LBD) | 139 |
| 9.11. Erkennungsleistungen nach Items (LBD) | 139 |
| 9.12. Subjektiver Schwierigkeitsgrad (RBD) | 141 |
| 9.13. Subjektives Sicherheitsgefühl (RBD) | 141 |
| 9.14. Erkennungsleistungen nach Items (RBD) | 142 |
| 9.15. Subjektiver Schwierigkeitsgrad (HC) | 143 |
| 9.16. Subjektives Sicherheitsgefühl (HC) | 143 |
| 9.17. Erkennungsleistungen nach Items (HC) | 144 |
| 9.18. Subjektiver Schwierigkeitsgrad (Gesamt) | 145 |
| 9.19. Subjektives Sicherheitsgefühl (Gesamt) | 145 |
| 9.20. Erkennungsleistungen nach Items (Gesamt) | 146 |
| 10.1. Stimmerkennungsmodell in Anlehnung an Bruce & Young (1986). . . | 151 |
| 10.2. Ausschnitt aus einem Testbogen des AKT | 154 |
| 10.3. Durchschnittliche zeitliche Dauer der Stimuli | 156 |
| 10.4. Leistung der Gesamtgruppe nach Stimulidauer | 157 |
| 10.5. Leistungsverlauf nach Stimulusdauer | 157 |
| 12.1. ANOVA zu den Variablen Alter und post-onset-time | 199 |
| 12.2. ANOVA zu den Variablen HADS-D und WMS-R | 199 |
| 12.3. ANOVA zu den Variablen des AKT | 199 |
| 12.4. ANOVA zur Variable <i>percentage correct (PC)</i> | 200 |
| 12.5. Deskriptive Statistik der Mittelwerte der Gruppen LBD und RBD . . | 200 |
| 12.6. U-Test: Erkennungsleistung der Gruppen LBD und RBD | 201 |
| 12.7. Deskriptive Statistik der Mittelwerte der Gruppen RBD und HC . . | 201 |
| 12.8. U-Test: Erkennungsleistung der Gruppen RBD und HC | 202 |

Abbildungsverzeichnis

12.9. Deskriptive Statistik der Mittelwerte der Gruppen LBD und HC . . . 202

12.10U-Test: Erkennungsleistung der Gruppen LBD und HC 202

12.11Mittelwertvergleich zwischen den Leistungen der Amnestischen und
der Wernicke-Aphasiker 203

12.12Deskriptive Statistik der Erkennungsleistung (*percentage correct [PC]*) 203

12.13Tests der Innersubjekteffekte 204

12.14Tests der Zwischensubjekteffekte 204

12.15U-Test: Erkennungsleistung auf CVC-Ebene der Gruppen LBD und HC 205

12.16U-Test: Erkennungsleistung auf W-Ebene der Gruppen LBD und RBD 205

12.17U-Test: Erkennungsleistung auf W-Ebene der Gruppen LBD und HC 206

12.18U-Test: Erkennungsleistung auf W-Ebene der Gruppen RBD und HC 206

12.19Korrelation zwischen Dauer und *percentage correct (PC)* 206

12.20*percentage correct (PC)* der zeitkritischen Stimuli 206

12.21Korrelationen zwischen den Vortests und der Erkennungsleistung (LBD) 207

12.22Korrelationen zwischen den Vortests und der Erkennungsleistung (RBD) 207

12.23Korrelationen zwischen den Vortests und der Erkennungsleistung (HC) 208

12.24Korrelationen zwischen den Vortests und der Erkennungsleistung (BD) 209

12.25Korrelationen zwischen den Vortests und der Erkennungsleistung (Ge-
samt) 210

12.26Korrelationen der Ratingskalen (LBD, V) 211

12.27Korrelationen der Ratingskalen (LBD, CVC) 211

12.28Korrelationen der Ratingskalen (LBD, W) 211

12.29Korrelationen der Ratingskalen (LBD, S) 212

12.30Korrelationen der Ratingskalen (RBD, V) 213

12.31Korrelationen der Ratingskalen (RBD, CVC) 213

12.32Korrelationen der Ratingskalen (RBD, W) 214

12.33Korrelationen der Ratingskalen (RBD, S) 214

12.34Korrelationen der Ratingskalen (HC, V) 215

12.35Korrelationen der Ratingskalen (HC, CVC) 215

12.36Korrelationen der Ratingskalen (HC, W) 216

12.37Korrelationen der Ratingskalen (HC, S) 216

12.38Korrelationen der Ratingskalen (Gesamt, V) 217

12.39Korrelationen der Ratingskalen (Gesamt, CVC) 217

12.40Korrelationen der Ratingskalen (Gesamt, W) 218

12.41Korrelationen der Ratingskalen (Gesamt, S) 218

13.1. Informationsblatt für die Distraktoren bzw. Angehörigen 220

Abbildungsverzeichnis

| | |
|--------------------------------------------------------------------|-----|
| 13.2. Informationsblatt für die Untersuchungsteilnehmer | 221 |
| 13.3. Anamnesebogen | 222 |
| 13.4. Testbogen "Verbale Paarerkenung 1" der WMS-R | 223 |
| 13.5. Auswertungsblatt A-K-T | 224 |
| 13.6. Probetestbogen A-K-T | 225 |
| 13.7. Testbogen Form A | 226 |
| 13.8. Testbogen Form A ₁ | 227 |
| 13.9. Testbogen Form A ₂ | 228 |
| 13.10 Testbogen der HADS-D, Subtest Depression (Seite 1) | 229 |
| 13.11 Testbogen der HADS-D, Subtest Depression (Seite 2) | 230 |
| 13.12 Bogen zur Erfassung der Wiedererkennungsleistung | 231 |
| 13.13 Bewertungsbogen (Seite 1) | 232 |
| 13.14 Bewertungsbogen (Seite 2) | 233 |

Tabellenverzeichnis

| | |
|----------------------------------------------------------------------------------------------------------|-----|
| 3.1. Vokale im deutschen Sprachsystem | 19 |
| 3.2. Plosive im deutschen Sprachsystem | 20 |
| 3.3. Frikative im deutschen Sprachsystem | 22 |
| 4.1. Intraindividuelle Unterschiede (Messergebnisse der Stimmenanalyse) | 38 |
| 4.2. Übersicht Grundfrequenzen | 40 |
| 5.1. Streuung der Wiedererkennungsrates gesunder Probanden in ausge- wählten Untersuchungen | 56 |
| 5.2. Übersicht über Stimmerkennungsstudien mit beeinträchtigten Pro- banden | 59 |
| 5.3. Leistungen der dyslektischen und der Kontrollprobanden | 60 |
| 5.4. Wiedererkennungsrates berühmter Stimmen (vor- vs. rückwärts) . . . | 65 |
| 5.5. Wiedererkennungsrates berühmter Stimmen (unterschiedliche Tempi) | 65 |
| 7.1. Leistungen im Bereich der Gesichter- bzw. Stimmentifizierung . . . | 89 |
| 7.2. Leistungen im Bereich der Gesichter- bzw. Stimmentifizierung . . . | 90 |
| 7.3. Ergebnisse der Sprechererkennungs- und Diskriminierungsaufgabe . . | 91 |
| 7.4. Ergebnisse der Stimmverarbeitungsaufgaben | 92 |
| 7.5. Ergebnisse der Stimmverarbeitungsaufgaben 2 | 93 |
| 8.1. Übersicht über Auswahl- und Ausschlusskriterien | 110 |
| 8.2. Set bei weiblichen Zielstimmen | 114 |
| 8.3. Set bei männlichen Zielstimmen | 115 |
| 8.4. Übersicht über die verwendeten Stimuli nach Stimulikategorie . . . | 116 |
| 9.1. Übersicht über Alter und <i>p.-o.-time</i> | 121 |
| 9.2. Ergebnisse in der WMS-R und dem HADS-D | 123 |
| 9.3. Prozentränge des AKT (sign.) | 125 |
| 9.4. Übersicht über die Gesamtleistung | 128 |

Tabellenverzeichnis

| | |
|----------------------------------------------------------------------------|-----|
| 9.5. Leistungen nach Stimuli | 131 |
| 9.6. Leistungen auf Vokalebene | 132 |
| 9.7. Leistungen auf CVC-Ebene | 132 |
| 9.8. Leistungen auf Wortebene | 133 |
| 9.9. Leistungen auf Satzebene | 134 |
| 9.10. Übersicht über die verwendeten Stimuli nach Stimulidauer | 135 |
| 9.11. Statistisch relevante Korrelationen (RBD) | 136 |
| 9.12. Statistisch relevante Korrelationen (BD) | 137 |
| 9.13. Ausgewählte Korrelationen der Ratingskalen (LBD) | 140 |
| 9.14. Ausgewählte Kovarianzen der Ratingskalen (RBD) | 142 |
| 9.15. Ausgewählte Kovarianzen der Ratingskalen (HC) | 144 |
| 9.16. Ausgewählte Kovarianzen der Ratingskalen (Gesamt) | 145 |
| 10.1. Einzelfalleleistungen im Vergleich zur RBD-Gesamtgruppe | 155 |
| 12.1. Deskriptive Statistik der Gruppe 1 | 190 |
| 12.2. Deskriptive Statistik der Gruppe 2 | 190 |
| 12.3. Deskriptive Statistik der Gruppe 3 | 191 |
| 12.4. Erreichte Rohwerte (HADS-D / WMS-R) bzw. Prozenträge (AKT) | 193 |
| 12.5. Akustische Parameter der Distraktorstimmen | 194 |
| 12.6. Schwierigkeitsgrad vs. Stimmerkennung (LBD) | 195 |
| 12.7. Sicherheitsgefühl vs. Stimmerkennung (LBD) | 195 |
| 12.8. Schwierigkeitsgrad vs. Stimmerkennung (RBD) | 196 |
| 12.9. Sicherheitsgefühl vs. Stimmerkennung (RBD) | 196 |
| 12.10. Schwierigkeitsgrad vs. Stimmerkennung (HC) | 197 |
| 12.11. Sicherheitsgefühl vs. Stimmerkennung (HC) | 197 |
| 12.12. Schwierigkeitsgrad vs. Stimmerkennung (Gesamt) | 198 |
| 12.13. Sicherheitsgefühl vs. Stimmerkennung (Gesamt) | 198 |
| 14.1. Technische Daten des verwendeten Mikrofonmoduls | 235 |
| 14.2. Technische Daten des verwendeten Speiseadapters | 235 |
| 14.3. Technische Daten des verwendeten DAT-Rekorders | 236 |

1. Einführung

1. Einführung

„An old friend telephones you unexpectedly and you know who it is before he declares himself. After watching a TV talk show for a few minutes, you know which panel member is speaking before the camera picks her up. You know that the voice behind you in the crowded elevator is unfamiliar, but you develop a clear notion of the age and sex of the speaker and some indication about the region of the country where the speaker was born“ (Bricker & Pruzansky 1974:295).

Wenn eine Person spricht, so produziert sie ein komplexes akustisches Signal, das in erster Linie einen semantischen Inhalt transportieren soll. Hörer, die mit der gesprochenen Sprache vertraut sind, können dem Sprecherschall diesen semantischen Inhalt entnehmen und in den meisten Fällen wiederholen oder zumindest sinngemäß wiedergeben. Wie die Beispiele von Bricker & Pruzansky (siehe oben) nahe legen, enthält das Signal neben den linguistischen Anteilen noch weitere, so genannte paralinguistische Elemente, die es dem Hörer ermöglichen zusätzliche Aussagen über den Sprecher zu treffen. So kann er z.B. in den meisten Fällen Angaben zu dem Geschlecht des Sprechers machen, zu dessen Alter, dessen emotionaler Verfassung, sogar zu dessen Gesundheitszustand (z.B. bei Erkältungen) und – mit Einschränkungen – zu dessen Bildungsstand.

Die menschliche Stimme transportiert also nicht nur Sprache, sie verleiht uns vielmehr ein „akustisches Gesicht“, das wichtige affektive und identitätsstiftende Informationen über uns beinhaltet und ausdrückt. Darüber hinaus ist unsere Stimme zweifelsfrei seit Jahrtausenden das wichtigste akustische Signal im Zusammenleben mit unseren Mitmenschen und unserer Interaktion miteinander und somit für unser Sozialleben von essenzieller Bedeutung.

Obwohl sich die im obigen Zitat geschilderten Situationen deutlich voneinander unterscheiden, so haben sie dennoch eine Gemeinsamkeit: in allen Fällen gelingt es einem Hörer, sprecherspezifische, paralinguistische Informationen aus einem Schallsignal zu extrahieren und zentralnervös zu verarbeiten. Ob dies auch Patienten mit einer links- bzw. rechtshemisphärischen Läsion gelingt, steht im Mittelpunkt dieser Arbeit. Dazu wurde ein quasi-experimentelles Design einer Stimmerkennungsaufgabe entwickelt und an aphasischen Patienten mit einer linkshemisphärischen Läsion, an nicht-aphasischen Patienten mit einer rechtshemisphärischen Läsion und an zerebral gesunden Kontrollprobanden durchgeführt. Das genaue methodische Vorgehen, die Ergebnisse, deren Interpretation und ein Ausblick auf noch offene Fragestellungen werden im **zweiten** Teil der Arbeit vorgestellt.

1. Einführung

Der **erste** Teil der Arbeit beginnt zunächst mit einer Einführung in die Grundlagen der Phonetik und Akustik. Ihr folgt eine ausführliche Übersicht über die unterschiedlichen Merkmale der Stimme, der Sprache und der Sprechweise sowie der dafür relevanten akustischen Parameter, die für das Erkennen eines Sprechers wesentlich sind. Kapitel 5 beleuchtet im Anschluss das Phänomen der Sprechererkennung aus zwei ganz unterschiedlichen Richtungen: zum einen wird die Sprechererkennung durch linguistisch naive Personen im *Alltag* vorgestellt, zum anderen die professionelle Sprechererkennung unter dem Blickwinkel der wissenschaftlichen Kriminalistik und *Forensik*. Im sechsten Kapitel erfolgt eine kurze Einführung in die (Neuro-) Anatomie und (Neuro-) Physiologie des physiologischen Hörens. Den Abschluss des ersten Teils bildet letztlich eine Einführung in die neurolinguistischen und -psychologischen Grundlagen mit den Schwerpunkten *Aphasiologie* und *Agnosien*.

Der Schwerpunkt des ersten Teils der Arbeit liegt infolgedessen auf den theoretischen Grundlagen der Sprechererkennung durch menschliche Hörer, die aus unterschiedlichen wissenschaftlichen Disziplinen stammen (z.B. Linguistik, Physik, Biologie, Neurologie etc.) und für das Verständnis der Herangehensweise im experimentellen Teil der Arbeit relevant sind.

Neben der Sprecher- bzw. Stimmerkennung (linguistisch naiver) Personen (*speaker recognition by listening*, *SRL* oder *casual speaker recognition* (Nolan 1990:457)) unterscheiden Bricker & Pruzansky (1974:295) zusätzlich die maschinelle oder automatisierte Sprechererkennung (*speaker recognition by machine*, *SRM* (bzw. *technical speaker recognition* bei Nolan 1990:458) sowie die Sprechererkennung durch die visuelle Analyse von Spektrogrammen (*speaker recognition by visual inspection of spectrograms*) (vgl. auch Hollien 1990:189).

Zwar haben die Forschungen v.a. im Bereich der automatisierten Sprechererkennung einen großen Beitrag zum Verstehen der SRL erbracht, dennoch spielt sie im Zusammenhang mit der Fragestellung dieser Arbeit eine untergeordnete Rolle und soll nicht näher vorgestellt werden. Die Gründe hierfür liegen v.a. darin, dass die SRM zum Identifizieren bzw. Verifizieren einer Sprecherstimme i.d.R. andere Parameter nutzt als dies bei menschlichen Hörern vermutlich der Fall ist. Eine ähnlich untergeordnete Rolle spielt auch die Sprechererkennung durch die visuelle Analyse von Spektrogrammen. Diese Methode findet nahezu ausschließlich im forensischen Bereich Verwendung und wird in Kapitel 5.2 kurz vorgestellt.

2. Begriffsbestimmung

2. Begriffsbestimmung

Da die Terminologie in den Veröffentlichungen zur Sprecher- oder Stimmerkennung oftmals mit unterschiedlicher Bedeutung verwandt wird, folgen in diesem Kapitel zunächst die Erklärungen und Definitionen der im Rahmen der vorliegenden Arbeit verwendeten Termini.

Hecker (vgl. 1971:2), Nolan (vgl. 1990:457), Van Dommelen (vgl. 1990:259) und zahlreiche andere Autoren verstehen unter *speaker recognition* (hier: Sprechererkennung) ganz allgemein den Entscheidungsprozess, der auf der Analyse sprecherabhängiger Stimmmerkmale beruht.

Darüber hinaus unterscheidet bereits Hecker (1971:2) zwei Unterbegriffe der Sprechererkennung, nämlich *identification* (*Identifizierung*) und *discrimination* (*Diskriminierung*). Bei der Identifizierung versucht ein Hörer anhand *einer* ausgewählten Stimmprobe den Sprecher zu identifizieren. Der Hörende muss dem Signal dazu die relevanten paralinguistischen Merkmale entnehmen und mit den bereits vorhandenen und abgespeicherten Referenz-Merkmalen einer endlichen Menge (nämlich der Menge der ihm bekannten Personen) vergleichen. Dabei bleibt jedoch unklar, ob der Hörer die Stimme tatsächlich identifiziert, sie also einem spezifischen Sprecher zuordnen kann, oder ob er lediglich eine bekannt / unbekannt-Entscheidung treffen muss. Dies ist insofern relevant, da sich diese beiden Leistungen qualitativ deutlich voneinander unterscheiden (vgl. Kapitel 7.3.2).

Unter Diskriminierung versteht Hecker (1971:2) die Leistung, nach dem Hören *zweier* Stimmproben eine Aussage darüber treffen zu können, ob es sich um ein und denselben oder um zwei unterschiedliche Sprecher handelt. Dazu muss der Hörer die paralinguistischen Informationen des Musters A zunächst extrahieren, verarbeiten und kurzzeitig speichern, um sie anschließend mit den paralinguistischen Parametern des Musters B vergleichen zu können. Die von ihm abgespeicherten Referenz-Muster bekannter bzw. vertrauter Personen spielen dabei keine Rolle.

Eine etwas andere Definition schlagen Van Lancker & Kreiman (1987) bzw. Van Lancker, Kreiman & Cummings (1989) vor. Sie verstehen unter *discrimination* (Diskriminieren) ebenfalls das Treffen einer gleich / ungleich-Entscheidung, allerdings nur bei *fremden* Stimmen. *Speaker recognition* (Sprechererkennung) „involves identification of a speaker known to the listener“ (Van Lancker, Kreiman & Cummings 1989:665). Auch hier bleibt die genaue Bedeutung von *identification* (Identifizierung) offen (siehe oben).

2. Begriffsbestimmung

Insbesondere auf Grund der unklaren Verwendung des Begriffes *Identifizierung*, sollen im Rahmen dieser Arbeit folgende Termini Verwendung finden:

- Unter **Stimm- oder Sprechererkennung** soll, wie von Van Lancker und anderen Autoren vorgeschlagen, die Leistung verstanden werden, anhand eines Stimmusters eine korrekte bekannt / unbekannt-Entscheidung treffen zu können.
- Unter **Sprecheridentifizierung** soll hier die Leistung verstanden werden, ein Stimmuster einer bekannten bzw. vertrauten spezifischen Person korrekt zuzuordnen. Diese Leistung kann z.B. durch das Nennen des Personennamens oder das Auswählen einer Fotografie überprüft werden (vgl. z.B. Van Lancker & Canter 1982, Van Lancker & Kreiman 1987, Nakamura et al. 2001:1048).
- Unter **Stimmdiskriminierung** wird die Leistung verstanden, anhand zweier (oder mehrerer) fremder Stimmuster eine gleich / ungleich-Entscheidung zu treffen.

Teil I.

Theoretische Grundlagen

3. Grundlagen der Phonetik und Akustik

3. Grundlagen der Phonetik und Akustik

Dass ein Hörer einen Schuss als Schuss, eine Glocke als Glocke oder ein Klavier als Klavier wahrnehmen kann, beruht auf der erfolgreichen Analyse und Verarbeitung unterschiedlicher Parameter physikalisch messbarer Schallschwingungen (z.B. Frequenzen, Amplituden etc.). Auch das Verstehen gesprochener Sprache und das Erkennen eines menschlichen Sprechers anhand seiner Stimme, z.B. durch die Analyse interindividueller Frequenzbereiche, beruht auf der Verarbeitung akustischer Phänomene. Dieses Kapitel soll daher zunächst in die für die Fragestellung relevanten physikalisch-akustischen Grundlagen der **Produktion** und **Rezeption** linguistischer und paralinguistischer Informationen einführen.

Grundlage der sprachlichen Kommunikation des Menschen ist die Verknüpfung und Aneinanderreihung unterschiedlicher Laute zu Lautketten, denen ein Hörer i.d.R. einen bestimmten semantischen Inhalt zuordnen kann. So steht beispielsweise die Lautkette [ˈfuːsbal] im Deutschen bzw. [ˈfʊtbɔːl] im Englischen (a) für ein rundes, hohles Spiel- oder Sportgerät aus Leder oder Kunststoff bzw. (b) für die dazugehörige Sportart. Damit eine hörende Person eine solche Lautkette korrekt interpretieren kann und um Missverständnisse zu vermeiden, ist es unverzichtbar, dass alle Angehörigen einer Sprachgemeinschaft die relevanten Laute auf sehr ähnliche Art und Weise produzieren. Besonders gravierend ist dies bei so genannten Minimalpaaren, also Wörtern oder Morphemen, die sich nur durch ein Phonem unterscheiden (z.B. [ˈtanə] vs. [ˈkanə]). Die sprachwissenschaftliche Disziplin, die in erster Linie solche lautlichen Anteile des menschlichen Kommunikationsprozesses zum Untersuchungsgegenstand hat, ist die *Phonetik*¹ (vgl. Kluge 2002:700, Pétursson & Neppert ³2002:14).

Die bedeutungsdifferenzierende Funktion der einzelnen Laute bleibt von der Phonetik unberührt und ist Gegenstand der *Phonologie*, auf die in diesem Rahmen jedoch nicht weiter eingegangen werden soll.

Traditionell unterscheidet die Linguistik drei Bereiche der Phonetik:

1. die **artikulatorische Phonetik** beschreibt die anatomischen und physiologischen Voraussetzungen und Gegebenheiten der Erzeugung von Sprachlauten durch einen Produzenten,

¹Bei dem Begriff *Phonetik* handelt es sich streng genommen um einen Neologismus der sich aus dem griechischen Wort *phōnē* (= Stimme, Ton, Laut, Rede, Sprache) und dem zugehörigen Adjektiv *phōnētikós* (= zur Stimme gehörig) zusammensetzt.

3. Grundlagen der Phonetik und Akustik

2. die **akustische Phonetik** beschreibt die Laute einer Sprache bezüglich ihrer physikalisch messbaren Substanz, also v.a. die Luftbewegungen und -schwingungen zwischen Produzent und Rezipient und
3. die **auditive Phonetik** beschreibt die Sprachlaute im Hinblick auf ihre Aufnahme und Verarbeitung durch einen Rezipienten.

3.1. Artikulatorische Phonetik

In der linguistischen und phonetischen Literatur wird der Begriff *Artikulation* mit unterschiedlichen Bedeutungen verwandt. Etymologisch wurzelt die Bezeichnung *artikulieren* (= [sorgfältig] aussprechen) in dem lateinischen Begriff *articulare* (= gliedern) und stellt eine Ableitung des Wortes *articulus* (= Abschnitt, Teil, Glied) dar (vgl. Kluge ²⁴2002:62), wodurch die Möglichkeit unterstrichen wird, Gesprochenes in unterschiedliche Einheiten zergliedern zu können (z.B. Worte, Sätze).

Unter Artikulation versteht man darüber hinaus recht allgemein die Tätigkeit aller am Sprechvorgang beteiligten Organe, also beispielsweise Mundlippen, Unterkiefer, Zunge etc. Streng genommen sind somit auch Atmung und Phonation Teilbereiche der Artikulation. Im Rahmen dieser Arbeit soll jedoch, wie in der linguistischen Literatur allgemein üblich, lediglich die Tätigkeit und das Zusammenwirken der supraglottalen Anteile unter dem Begriff Artikulation subsumiert werden (vgl. Pétursson & Neppert ³2002:45).

Obwohl Atmung und Phonation im Allgemeinen nicht Teil der unmittelbaren Artikulation sind, gehören sie dennoch zu den drei unabdingbaren Funktionskreisen, die für eine Sprachproduktion unerlässlich sind.

3.1.1. Atmung (Initiator)

Ogleich die Atmung in mehrfacher Hinsicht eine wesentliche Voraussetzung für den erfolgreichen Phonationsvorgang darstellt, soll an dieser Stelle nur ein sehr kurzer Überblick über ihren befundfreien und ungestörten Ablauf erfolgen.

Die Einatmungsmuskeln bewirken die Anhebung des Brustkorbes und das Absenken des Zwerchfells. Dadurch kommt es zu einer Vergrößerung des inneren Lungenvolumens wodurch gegenüber der Außenwelt des Körpers ein Unterdruck entsteht. Die Folge ist das Einströmen der Einatemluft (*Inspiration*) durch die Luftröhre (*Trachea*) in die Bronchien, die Bronchiolen und zuletzt in die Lungenbläschen (*Alveolen*), die von einem Netz von Blutgefäßen (*Lungenkapillaren*) umgeben werden (vgl. Faller & Schünke ¹⁴2004:387). Hier findet dann der auf dem Prozess der Diffu-

3. Grundlagen der Phonetik und Akustik

sion beruhende Gasaustausch zwischen Luft und Blut statt (vgl. Faller & Schünke¹⁴2004:393).

Für die eigentliche Phonation, also die Stimmtoneerzeugung, ist die Ausatmung (*Exspiration*) von besonderer Bedeutung. Die Ausatemungsmuskeln bewirken dabei eine Verkleinerung der Brusthöhle. Dadurch verringert sich das innere Lungenvolumen, was zu einer Gasdruckerhöhung führt (vgl. Ladefoged⁵2006:133f, Faller & Schünke¹⁴2004:405). Die Druckdifferenz ermöglicht das Ausströmen der Luft, die als Energiequelle für die eigentliche Schallerzeugung genutzt wird (vgl. Pétursson & Neppert³2002:48).

3.1.2. Phonation (Generator)

Der Kehlkopf (*Larynx*) (siehe Abbildung 3.1 auf der nächsten Seite) bildet bei allen landbewohnenden Wirbeltieren – und somit auch beim Menschen – den Eingang in die Luftröhre (*Trachea*) und verursacht bei der Exspiration die erste Modifikation des pulmonalen Luftstroms. Der *Larynx* ist von einem Knorpelgerüst umgeben, das sich aus den Resten der letzten Kiemenbogen entwickelt hat.

Beim Menschen besteht das laryngeale Knorpelgerüst aus dem Schildknorpel (*Cartilago thyreoides*, Thyroid), dem siegelringförmigen Ringknorpel (*Cartilago cricoidea*, Cricoid) und den paarigen, mit ihm gelenkig verbundenen Stell- oder Aryknorpeln (*Cartilagine arytaenoideae*, Arytenoid). Rostral ist der Larynx durch den Kehlkopfdeckel (*Epiglottis*) verschließbar, damit z.B. der Nahrungsbrei über die heruntergezogene *Epiglottis* in die Speiseröhre (*Ösophagus*) geleitet werden kann. Zwischen den Aryknorpeln und der Innenwand der Schildknorpel spannen sich die Stimmbänder (*Ligamentum vocale*) und Stimmlippenmuskeln (*Musculus vocalis*), die zwischen sich die Stimmritze (*Glottis*) freilassen (vgl. Faller & Schünke¹⁴2004:375f). Diese letztgenannten Strukturen (Epiglottis, Stimmbänder und -lippen) erfüllen eine so genannte Sphinkterfunktion und schützen die Luftwege vor dem Eindringen von Fremdkörpern, wie z.B. Lebensmitteln und Flüssigkeiten. Dies ist die Primärfunktion des Larynx. Die für die Phonation relevante Stimmbildung und Lauterzeugung stellt lediglich eine Sekundärfunktion dar.

Die eigentliche Stimmtoneerzeugung im Larynx, die für die Artikulation aller Vokale² und stimmhafter Konsonanten unverzichtbar ist, beginnt mit dem Verengen bzw.

²Diese Bezeichnung geht auf Pike (1943) zurück, der „für die zentralen oralen friktionslosen (...) Laute die Bezeichnung *Vokale*, für alle anderen *Konsonanten* vorgeschlagen [hat], um so eine phonetische Klassifikation von einer phonologischen in Vokale und Konsonanten zu trennen“ (Kohler²1995:61f, Sperrungen im Original).

3. Grundlagen der Phonetik und Akustik

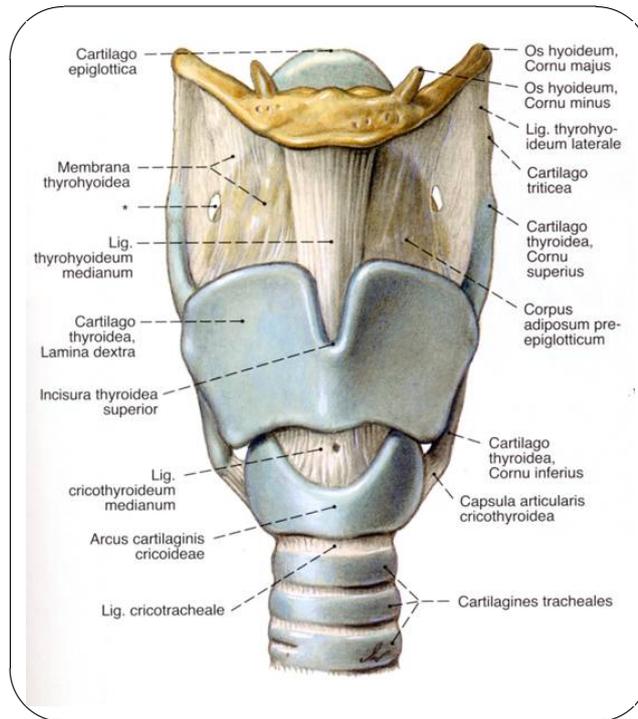


Abbildung 3.1.: Kehlkopf von ventral
(aus: Putz & Pabst ²¹2004:125)

Verschließen der Glottis (vgl. Abbildung 3.2). Die Aryknorpel dienen dabei als Ad-
duktoren der Stimmlippen.

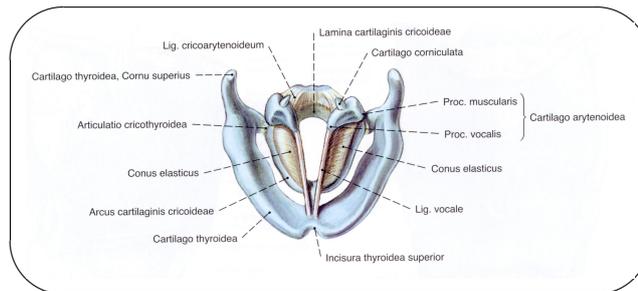


Abbildung 3.2.: Kehlkopfknorpel mit Glottis von oben ventral
(aus: Putz & Pabst ²¹2004:126)

Bei der Phonation mit **festem Stimmeinsatz** sind die Stimmlippen zunächst geschlossen (vgl. Phase a_1 in Abbildung 3.3 auf der nächsten Seite). Durch die Aktivität der Ausatemungsmuskulatur kommt es zu einer Erhöhung des supraglottalen Luftdruckes, der, wenn er groß genug ist, den Glottisverschluss aufsprengt (vgl.

3. Grundlagen der Phonetik und Akustik

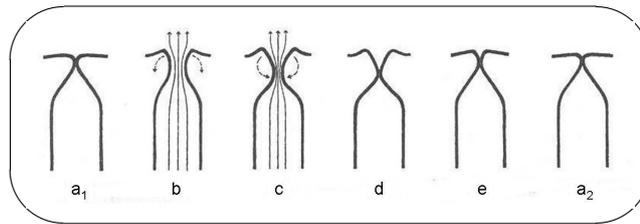


Abbildung 3.3.: Sog- bzw. Bernoulli-Effekt
(verändert und ergänzt nach Neppert ⁴1999:120)

Phase *b*). Innerhalb der entstandenen Düse, bestehend aus Stimmlippen und *Conus elasticus*, strömt die Luft mit hoher Geschwindigkeit in die sich anschließenden supralaryngealen Hohlräume, die zusammengefasst als *Ansatzrohr* bezeichnet werden. Auf Grund der großen Strömungsgeschwindigkeit innerhalb der Düse gegenüber der Umgebung entsteht ein statischer Unterdruck. Dieses Phänomen beruht auf den strömungsdynamischen Gesetzen, die der Mathematiker und Physiker Daniel BERNOULLI (1700-1782) im Rahmen der *Bernoullischen Strömungsgleichung* beschrieben hat.

Im weiteren Verlauf führen die Sogkräfte des Unterdruckes die Glottis beginnend im Bereich des *Conus elasticus* wieder zusammen (vgl. Phase *c*, *d* und *e*). Dabei wirken auch die so genannten *myoelastischen Rückstellkräfte* mit, die auf dem muskulären Grundtonus der Stimmlippen beruhen. Ist der Zyklus des Verschließvorganges abgeschlossen, kann der nächste Öffnungszyklus beginnen (vgl. Phase *a*₂).

Bei der Phonation mit **weichem** oder **behauchtem Stimminsatz** laufen die Vorgänge sehr ähnlich ab. Hier beginnt der Öffnungszyklus jedoch nicht mit einem Aufsprengen der Glottis durch eine Erhöhung des supraglottalen Druckes, sondern mit einer Adduktion der Stimmlippen (vgl. Phase *b*). Während des sukzessiven Verschließens der Glottis nimmt die Strömungsgeschwindigkeit der ausströmenden, pulmonal-exhalatorischen (Ausatem-) Luft zu (siehe oben) und es kommt zu dem bereits erwähnten Unterdruck zwischen der entstehenden Düse und ihrer Umgebung (vgl. Phase *c*). Gemeinsam mit den myoelastischen Rückstellkräften stellt dieser Unterdruck nun die Kraft zur Verfügung, die zu einem Verschluss der Düse führt und die Ausgangssituation für den nächsten Öffnungszyklus darstellt (vgl. Phase *d*, *e* und *a*₂) (vgl. Neppert ⁴1999:122).

Zwar werden die akustischen Folgen dieses Prozesses erst in Kapitel 3.2 näher erläutert, jedoch soll bereits an dieser Stelle darauf hingewiesen werden, dass nicht die Schwingung der Stimmlippen als Gesamtkörper das akustische Signal erzeugt, wie es z.B. bei einer Stimmgabel der Fall ist, sondern dass es vielmehr durch die auf Grund des oben beschriebenen Bernoulli-Effekts verursachten impulsartigen Störungen des

3. Grundlagen der Phonetik und Akustik

Luftdrucks unmittelbar oberhalb der Glottis (so genannte *Kavitätenbildung*) entsteht (vgl. Kuttruff 2004:3, Pompino-Marschall ²2003:88, 100)³.

Diese Druckimpulse stellen ein noch völlig unspezifisches Signal dar, das erst dann zu einem Sprachlaut wird, wenn sein zeitlicher Verlauf, also sein Spektrum, modifiziert wird. Je nach Ort der Enge und Modifikation im Ansatzrohr zeigt der Sprachlaut dann seine spezifische spektrale Charakteristik. Wie dies geschieht soll im folgenden Kapitel anhand verschiedener Laute des deutschen Sprachsystems erläutert werden.

3.1.3. Artikulation (Modifikator)

Eine indirekte Modifikation erhält der Rohschall zunächst durch das Ansatzrohr (so genannte *Rohschallmodifikation*). Es umfasst alle supralaryngealen Hohlräume und erstreckt sich vom *Larynx* bis zu den Nasenmuscheln (*Conchae*). Für eine bessere Übersicht lässt sich das Ansatzrohr in unterschiedliche Räume gliedern. Der Abschnitt vom Schlund (*Pharynx*) bis zum weichen Gaumensegel (*Velum*) wird als **Pharynxraum** bezeichnet und besteht wiederum aus drei Teilen: dem *Laryngopharynx* (Stimm lippen bis Zungenbein, Nr. 1a), dem *Oropharynx* (Zungenbein bis *Velum*, Nr. 1b) und dem *Nasopharynx* (*Velum* bis *Choanen*, Nr. 1c). Der zweite große Hohlraum, der **Mundhohlraum** (Nr. 2), erstreckt von dem vorderen Gaumenbogen bis zu den Mundlippen. Streng physikalisch betrachtet handelt es sich bei diesem Raum nicht um einen einfachen Resonanzraum, sondern um zahlreiche kleinere Hohlräume, z.B. zwischen den Zähnen und der Wangeninnenseite. Der Einfachheit halber wird er jedoch als einzelner Raum angenommen. Dies gilt auch für den **Nasenhohlraum** (Nr. 3), der durch die Nasenscheidewand (*Septum nasi*) in zwei Räume geteilt wird und mit den bereits erwähnten *Conchae* ausgefüllt ist. Auch er besteht somit aus zahlreichen kleineren Hohlräumen (vgl. Abbildung 3.4 auf der nächsten Seite).

Da die Individualität einer Sprecherstimme ganz wesentlich von den anatomischen Bedingungen der supralaryngealen Hohlräume und deren Resonanzverhalten beeinflusst wird, soll an dieser Stelle die physikalische Entstehung von Resonanzen beispielhaft vereinfachend an dem Reflexionsverhalten eines neutralen Rohres veranschaulicht werden. Dieses 'künstliche' Ansatzrohr soll eine Länge von 17 cm und einen kreisrunden Durchmesser von 4 bis 5 cm aufweisen. Es zeige in seinem Gesamtverlauf [z] eine konstante Querschnittsfläche [σ] und besitze schallharte Wände, die für eine Totalreflexion sorgen. An seinem unteren Ende soll sich ein ebenso schallharter

³Dies soll jedoch nicht darüber hinwegtäuschen, dass im Tierreich durch das Aneinanderreiben von Körperteilen (Einsatz von Schrillkanten bzw. -flächen) Geräusche zur Informationsweitergabe eingesetzt werden können (so genannte *Stridulation*), wie z.B. bei Ameisen. Dies ist jedoch nicht mit der Funktionsweise der menschlichen Stimm lippen zu verwechseln.

3. Grundlagen der Phonetik und Akustik

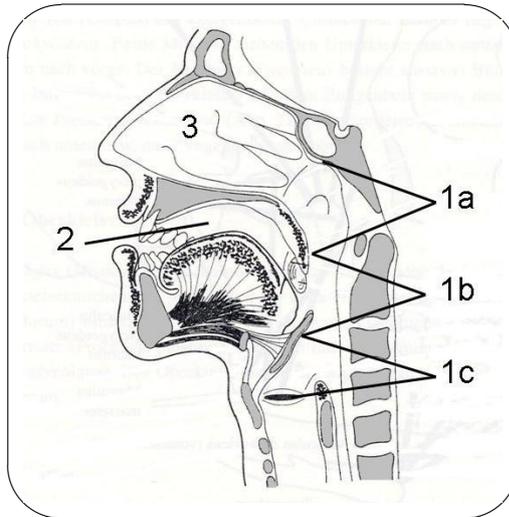


Abbildung 3.4.: Hohlräume im Ansatzrohr
(verändert und ergänzt nach Pétursson & Neppert ³2002:77)

Verschluss (entspricht der Glottis), an seinem oberen Ende eine Öffnung (entspricht den Mundlippen) befinden. Kommt es innerhalb dieses Rohres zu einem Impuls, also einer Luftdruckstörung, so werden die Luftmoleküle an der Innenseite des Rohres und an seinem verschlossenen Ende total reflektiert, während sie am offenen Rohrende auf keinen Reflexionswiderstand stoßen. Während der Schalldruckwechsel am geschlossenen Ende maximal [$p = \max$] ist, findet sich am offenen Ende stets der Luftdruck der äußeren Umwelt [$p = 0$]. Diese Randbedingungen ermöglichen das Entstehen so genannter *stehender Wellen*, wie sie – in Extremform – in Abbildung 3.5 auf der nächsten Seite dargestellt sind.

Die erste stehende Welle, die den entsprechenden Randbedingungen genügt, ist ein Cosinus mit der Wellenlänge λ von der vierfachen Rohrlänge ($\lambda_1 = 4 \times L$). Die zweite Resonanzfrequenz besitzt eine Wellenlänge von vier Drittel, die dritte von vier Fünftel der Rohrlänge ($\lambda_2 = \frac{4}{3} \times L$ bzw. $\lambda_3 = \frac{4}{5} \times L$) (vgl. Pompino-Marschall ²2003:107). Die Resonanzfrequenzen eines solchen theoretischen, neutralen Rohres lassen sich mit folgender Formel leicht berechnen (vgl. Pétursson & Neppert ³2002:135):

$$\text{Frequenz} = \frac{\text{Schallgeschwindigkeit (Luft)}}{\text{Wellenlaenge}}$$

oder

$$f_i = \frac{c_L}{\lambda_i}. \quad (3.1)$$

3. Grundlagen der Phonetik und Akustik

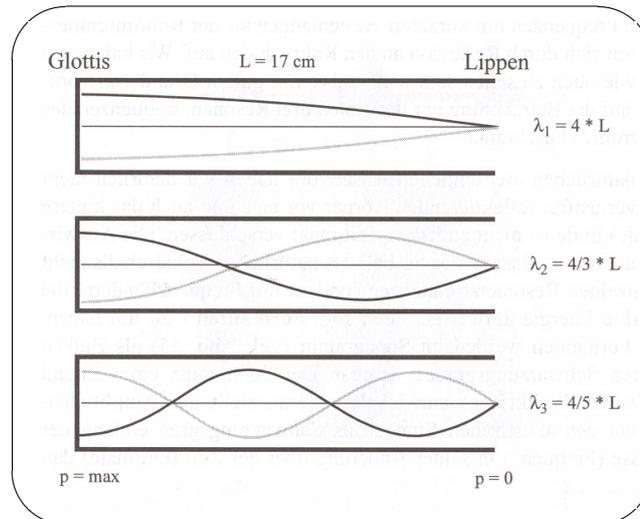


Abbildung 3.5.: Stehende Schalldruckwellen im neutralen Rohr
(aus: Pompino-Marschall ²2003:107)

Berücksichtigt man nun die angenommenen Maße des Rohres, so führt dies zu folgenden Resonanzfrequenzen:

$$F_1 = \frac{340}{4 \times 0,17} = \frac{340}{0,68} = 500 \text{ Hz},$$

$$F_2 = \frac{3 \times 340}{4 \times 0,17} = \frac{1020}{0,68} = 1500 \text{ Hz},$$

$$F_3 = \frac{5 \times 340}{4 \times 0,17} = \frac{1700}{0,68} = 2500 \text{ Hz}.$$

Ein solches neutrales Rohr zeigt also Resonanzfrequenzen im Verhältnis 1:3:5 (vgl. Pompino-Marschall ²2003:108).

Besonderheiten für die Sprechschallakustik

Betrachtet man das menschliche Ansatzrohr, so handelt es sich dabei selbstverständlich nicht um ein total reflektierendes Rohr wie es oben vorgestellt wurde. Im Gegenteil: es besteht aus mehreren in Form und Gestalt recht unterschiedlichen aneinandergereihten Hohlräumen, deren Wände mit weichem Gewebe bedeckt sind und somit nicht verlustfrei reflektieren. Dies zeigt sich im akustischen Ergebnis darin, dass beim Stimmschall keine einzelnen, klar messbaren Resonanzfrequenzen erkennbar sind, sondern relativ breite Frequenzbänder und -maxima, die *Formanten* genannt werden. Dabei ist der n-te Formant im Sprechschall das Äquivalent zur n-ten Resonanzfrequenz des Ansatzrohres (vgl. Neppert ⁴1999:48, Künzel 1987:26). Da die Lage

3. Grundlagen der Phonetik und Akustik

der Formanten durch das Resonanzverhalten der Hohlräume des Ansatzrohres beeinflusst werden, ist ihre Verteilung interindividuell unterschiedlich, obgleich sie sich selbstverständlich bei allen Sprechern in einem ungefähren Wertebereich bewegen (vgl. z.B. Hollien 1990:196). Dabei ist jedoch nicht die absolute Lage der Formanten auf der Frequenzskala entscheidend, sondern die Relationen der Formantpositionen. Dies erklärt, dass Vokoide trotz unterschiedlicher Formantpositionen in der Regel sprecherunabhängig erkannt werden können (Mann vs. Frau vs. Kind) (vgl. Neppert⁴1999:141). Dabei reichen bereits die beiden ersten Formanten (F_1 und F_2) aus, um einen Laut hinreichend zu charakterisieren.

Die für den eigentlichen Sprechakt entscheidende Modifikation erfährt der Rohschall durch die Tätigkeit und das Zusammenwirken unterschiedlicher supraglottaler Anteile. Dabei ist die Fülle der artikulatorischen Möglichkeiten und die Lautgestalt der verschiedenen Sprachen immens. Um alle Laute der unterschiedlichen Sprachen graphematisch adäquat und eindeutig darstellen zu können, reichen die Grapheme des lateinischen oder griechischen Alphabetes bei weitem nicht aus. Aus diesem Grunde stellt der Internationale Phonetikerverband (API bzw. IPA⁴) ein phonetisches Alphabet (IPA⁵) zur Verfügung. Bei den verwendeten Graphemen handelt es sich zum einen um abgewandelte und ergänzte Zeichen des griechischen und lateinischen Alphabetes, zum anderen um Neuschöpfungen. Im Rahmen dieses Unterkapitels sollen zunächst die Vokoide, im Anschluss daran die Kontoide geordnet nach ihrer Artikulationsart (z.B. Plosive, Nasale etc.) erläutert werden (vgl. Abbildung 3.8 auf Seite 23). Da an dieser Stelle lediglich auf die für das Deutsche relevanten Laute eingegangen werden soll, können die artikulatorischen Besonderheiten von Taps, Flaps, lateralen Frikativen, Clicks und Ejektiven bzw. Implosiven im Abschnitt *Weitere Kontoide* (ab Seite 23) nur ansatzweise erläutert werden.

Vokoide

Wie bereits an anderer Stelle angedeutet, handelt es sich bei dem mit Hilfe der Stimmlippen produzierten Rohschall um ein unspezifisches akustisches Signal, das erst durch eine charakterisierende Modifikation zu einem spezifischen Sprachlaut wird. Bei der Rohschallmodifikation zur Bildung von Vokoiden ist das Ansatzrohr stets geöffnet und lediglich die konvex gewölbte Zunge verändert, je nach zu artikulierendem Laut, ihre Lage.

⁴Association Phonétique Internationale bzw. International Phonetic Association.

⁵International Phonetic Alphabet.

3. Grundlagen der Phonetik und Akustik

Zu den artikulatorischen Hauptparametern der Vokoiden gehören:

1. die Zungenhöhe mit den Parametern

- hoch [i], [u],
- halbhoch [ɪ], [ʊ],
- obermittelhoch [e], [o],
- mittel [ə],
- untermittelhoch [ɛ], [ɔ],
- halbtief [æ], [ɶ],
- tief [a], [ɑ],

2. die Zungenlage mit den Parametern

- palatal [i] bis [a],
- zentral [ɪ] bis [ɶ],
- velar [u] bis [ɑ] und

3. die Lippenrundung mit den Parametern

- ungerundet (Kardinalvokale [i] bis [ɑ]) und
- gerundet (Kardinalvokale [ɔ] bis [u]) (verändert nach Pompino-Marschall ²2003:221; vgl. Kohler ²1995:66ff; Pullum & Ladusaw ²1996).

Als Referenzsystem für die Transkription von Vokoiden stützt sich das IPA auf das System der *Kardinalvokale*, das der Phonetiker Daniel JONES 1917 entwickelt hat, um die Vokalsysteme unterschiedlicher Sprachen miteinander vergleichen zu können. Ausgehend von dem höchsten und vordersten Vokoid [i] und dem tiefsten hintersten [ɑ], legte Jones konkrete Positionen der sprachenunabhängigen Vokoidrealisierungen fest und definierte damit die so genannten *primären Kardinalvokale* [i], [e], [ɛ], [a], [u], [o], [ɔ] und [ɑ] (vgl. Ladefoged ⁵2006:211-216, Pompino-Marschall ²2003:222, Clark & Yallop ²1995:22, Kohler ²1995:67).

Neben den primären Kardinalvokalen existieren darüber hinaus auch *sekundäre Kardinalvokale* (vgl. Abbildung 3.6), deren Systematisierung ebenfalls auf JONES zurückgeht (1956). Bei ihrer Bildung bleiben die Zungenstellungen erhalten und die Lippenrundungen werden vertauscht. So entstehen die Vokoiden [y], [ø], [œ], [ɕ], [ɸ], [ʌ], [ʏ]⁶, [ʉ]⁷, [ɨ]⁸ und [ɥ]⁹ (vgl. u.a. Ladefoged ⁵2006:216ff).

⁶So z.B. im Vietnamesischen (vgl. Pompino-Marschall ²2003:227).

⁷So z.B. im Türkischen und Japanischen (vgl. Pullum & Ladusaw ²1996:114).

⁸So z.B. im Russischen (vgl. Clark & Yallop ²1995:30).

⁹So z.B. im Schottischen (vgl. Clark & Yallop ²1995:28).

3. Grundlagen der Phonetik und Akustik

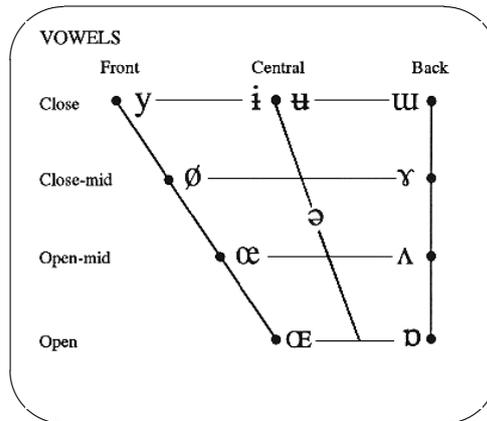


Abbildung 3.6.: Sekundäre Kardinalvokale

(Internetquelle: <http://www2.arts.gla.ac.uk/IPA/images/vowels.gif> (31.10.2005, verändert))

Zusätzlich zu den sekundären Kardinalvokalen empfiehlt die IPA weitere Transkriptionszeichen, nämlich die halbhohen Vokale [ɪ] bzw. [ʏ] (vorne, ungerundet bzw. gerundet), [ʊ] (hinten, gerundet), [æ] (vorne, halbtief), [ə] bzw. [ɐ] (obermittelhoch, ungerundet bzw. gerundet), [ɜ] bzw. [ɝ] (untermittelhoch, ungerundet bzw. gerundet) und [ɶ] (halbtief, ungerundet) (vgl. Pompino-Marschall ²2003:222f). Tabelle 3.1 gibt eine Übersicht über die im Deutschen auftretenden Vokale, das vollständige Vokaltrapez zeigt Abbildung 3.7 auf der nächsten Seite.

| front | central | back |
|----------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| [i] I gel, b ieten, sie | [y] Ü bel, h üten, fr üh | [u] U hr, s puten, U hu |
| [ɪ] i ch, b itten | [ʏ] ü ppig, H ütten | [ʊ] u m, B utter |
| [e] E sel, b eten, T ee | [ø] Ö fen, G oethe | [o] O fen, b oten, f roh |
| [ɛ] E cke, B etten | [œ] ö ffnen, G ötter | [ɔ] o ffen |
| [ɛ:] ä sen, b äten | | |
| | [a:] a ber, b aten, s ah | |
| | [a] A cker, h atten | |
| | [ə] B eute | |

Tabelle 3.1.: Vokale im deutschen Sprachsystem
(vgl. Pompino-Marschall ²2003:267)

3. Grundlagen der Phonetik und Akustik

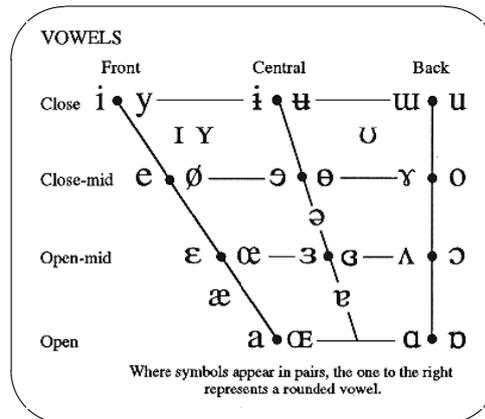


Abbildung 3.7.: Vokaltrapez

(Internetquelle: <http://www2.arts.gla.ac.uk/IPA/images/vowels.gif> (31.10.2005))

Plosive

Der Artikulationsmodus der *Plosive* (Verschlusslaut, Explosiv, Okklusiv) ist der vorübergehende totale Verschluss der Stimm- oder Mundlippen sowie des Nasenraumes. Die Bewegung vom offenen Ansatzrohr zum Verschluss zeigt akustisch *implosive Transitionen* (Formantbewegungen) und eine deutliche Reduktion der Amplitude. Bei stimmhaften Plosiven hält der Stimmtton – u.U. auch nur vorübergehend – an. Charakteristisch für das Wahrnehmen und Erkennen eines Plosivs ist das Geräusch während der Verschlussöffnung (*burst*) und die erneuten Formantbewegungen (*explosive Transitionen*) beim Übergang vom geschlossenen zum geöffneten Ansatzrohr (z.B. aufgrund der artikulatorischen Bewegung zum Folgevokal). Tabelle 3.2 gibt eine Übersicht über die im deutschen Lautsystem auftretenden Plosive.

| stimmhaft | stimmlos |
|------------------------------------------------|---------------------------------|
| [p] P ass, L ippe, l ump | [b] B ass, E bene |
| [t] T asse, L eute, l aut | [d] e del |
| [k] K asse, L ake, S ack | [g] L age |

Tabelle 3.2.: Plosive im deutschen Sprachsystem
(vgl. Pompino-Marschall ²2003:267)

3. Grundlagen der Phonetik und Akustik

Nasale

Auch *Nasale* erfordern, wie Plosive, einen kompletten oralen Verschluss. Anders als bei den Plosiven wird durch das Absenken des Gaumensegels der Nasenhohlraum als zusätzlicher Resonanzraum genutzt. Durch die Bildung eines oralen Verschlusses kommt es auch hier zu Formantbewegungen. Das erforderliche Absenken des Velums zeigt im Spektrum eine deutlich abgrenzbare Phase mit niedrigen Frequenzkomponenten. Anders als bei den Plosiven kann bei der Artikulation der Nasale der pulmonale Luftstrom auch während des oralen Verschlusses entweichen und es entsteht kein spezifisches Geräusch unter der oralen Verschlussöffnung. Eine abrupte spektrale Änderung zeigt sich jedoch bei der Anhebung des Velums, was zu einer Beendigung des Nasals führt. Für die Nasale sind im Deutschen folgende Artikulationsstellen möglich: labial [m] (z.B. in **Maß**, **Eimer** und **Leim**, labiodental¹⁰ [ɱ]), alveolar [n] (z.B. in **Nase**, **ohne** und **Sohn**) und velar [ŋ] (z.B. in **lange** und **Gang**).

Vibranten

Kurze intermittierende orale Verschlüsse erzeugen so genannte *Vibranten* oder *Trills*. Realisiert werden sie durch den leichten Verschluss zwischen einem elastischen Artikulator und der entsprechenden Artikulationsstelle. Der Verschluss wird durch einen pulmonalen Luftstrom gesprengt, der an der entstandenen Düse den bereits erläuterten Bernoulli-Effekt auslöst. Da lediglich die Mundlippen, die Zunge und das Zäpfchen die hierfür erforderliche Elastizität aufweisen, ist die Anzahl der möglichen Konstellationen der Artikulatoren sehr eingeschränkt. So treten im deutschen Sprachsystem lediglich die beiden folgenden Vibranten auf: das gerollte, alveolar gebildete Zungenspitzen-r [r] ist Bestandteil der deutschen Bühnensprache und v.a. im süddeutschen Sprachraum weit verbreitet, wohingegen das gerollte, uvular gebildete Zäpfchen-r [ʀ] eher im norddeutschen Sprachraum auftritt (vgl. Pompino-Marschall ²2003:195).

Frikative

Die *Frikative* (Spirant, Reibelaut, Engelaute) bilden die Lautklasse, die die artikulatorisch größte Ausdifferenzierung erlaubt. Um einen Frikativ zu artikulieren muss, anders als bei den lateralen Frikativen (vgl. Abschnitt *Weitere Kontoide* ab Seite 23), eine zentrale Enge gebildet werden, durch die die hindurchströmende Luft ein Geräusch verursacht. Je nach Lage der Enge im Ansatzrohr und ihrer geometrischen

¹⁰Dieser labiodentale Nasal findet sich aufgrund v.a. dialektal bedingter assimilatorischer Vorgänge auch im Deutschen (z.B. ⟨*fünf*⟩ → [fʏɱf]).

3. Grundlagen der Phonetik und Akustik

Gestalt variiert das akustische Signal. Dabei gilt: je länger das Ansatzrohr hinter der zu passierenden Enge ist, desto tiefer liegt die unterste Frequenz im entsprechenden Signal; im Deutschen ist sie somit beim [s] am höchsten und beim [x] bzw. [χ] am tiefsten. Tabelle 3.3 zeigt eine Übersicht der im Deutschen auftretenden Frikative.

| stimmhaft | stimmlos |
|-----------------------------------|--------------------------------|
| [f] F ass, laufen, Ruf | [v] W asser, ewig |
| [s] reißen, Reis | [z] S onne, reisen |
| [ʃ] s chon, Masche, rasch | [ʒ] G enie, Garage |
| [ç] C hemie, stechen, dich | |
| [x] s uchen, B uch | |
| [χ] m achen, D ach | [ʁ] R ast, E hre |
| [h] h asse, A horn | |

Tabelle 3.3.: Frikative im deutschen Sprachsystem
(vgl. Pompino-Marschall ²2003:267)

Approximanten

Da die *Approximanten* eine ähnliche Artikulation wie Vokale verlangen, werden sie gelegentlich auch als *Halbvokale* oder in der angloamerikanischen Literatur als *semi-vowels* bezeichnet (vgl. z.B. Clark & Yallop ²1995:44). Auch für ihre Produktion ist eine Enge erforderlich, die jedoch größer sein muss, so dass nicht das typische Frikativrauschen entstehen kann. Das deutsche Lautsystem zeigt lediglich einen solchen Approximanten, nämlich das palatal gebildete [j] in **ja** und **Ajax** (vgl. Ladefoged ⁵2006:15, Pompino-Marschall ²2003:266).

Laterale Approximanten

Die lateralen Approximanten, die abgekürzt einfach nur *Laterale* genannt werden, zeichnen sich durch einen zentralen Verschluss des Ansatzrohres aus, während eine seitliche Enge gebildet wird, die nicht Geräusch bildend ist. Der einzige laterale Approximant des deutschen Lautsystems ist das alveolar gebildete [l], das z.B. in **Lasso**, **Beule** und **Kohl** auftritt.

3. Grundlagen der Phonetik und Akustik

THE INTERNATIONAL PHONETIC ALPHABET (revised to 1993)
CONSONANTS (PULMONIC)

| | Bilabial | Labiodental | Dental | Alveolar | Postalveolar | Retroflex | Palatal | Velar | Uvular | Pharyngeal | Glottal |
|---------------------|----------|-------------|--------|----------|--------------|-----------|---------|-------|--------|------------|---------|
| Plosive | p b | | | t d | | ʈ ɖ | c ɟ | k ɡ | q ɢ | | ʔ |
| Nasal | m | ɱ | | n | | ɳ | ɲ | ŋ | ɴ | | |
| Trill | ʙ | | | r | | | | | ʀ | | |
| Tap or Flap | | | | ɾ | | ɽ | | | | | |
| Fricative | ɸ β | f v | θ ð | s z | ʃ ʒ | ʂ ʐ | ç ʝ | x ɣ | χ ʁ | ħ ʕ | h ɦ |
| Lateral fricative | | | | ɬ ɮ | | | | | | | |
| Approximant | | ʋ | | ɹ | | ɻ | j | ɰ | | | |
| Lateral approximant | | | | l | | ɭ | ʎ | ʟ | | | |

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

Abbildung 3.8.: Pulmonale Kontoide

(Internetquelle: <http://www2.arts.gla.ac.uk/IPA/pulmonic.html> (01.11.2005))

Weitere Kontoide

Getippte und geschlagene Laute, so genannte **Taps** und **Flaps**, kommen im Standarddeutschen nicht vor. Bei diesen Lauten ist zunächst ein kompletter, jedoch nur sehr kurzzeitiger, oraler und nasaler Verschluss erforderlich. Er wird durch Antippen oder Schlagen der Zungenspitze an eine Artikulationsstelle gebildet. Taps und Flaps treten an alveolarer [ɾ]¹¹ bzw. benachbarter Artikulationsstelle oder als Retroflexe [ɽ]¹² auf. Die **lateralen Frikative** findet man v.a. in afrikanischen Sprachen, wie z.B. dem Margi, in nordamerikanischen Indianersprachen, im Walisischen und auch im grönländischen Inuktitut (vgl. Pompino-Marschall ²2003:203, Pullum & Ladusaw ²1996:105, 107). Anders als bei den „zentralen“ Frikativen liegt bei den lateralen Frikativen die notwendige Enge nicht im „Längszentrum des Ansatzrohres“ (Pompino-Marschall ²2003:202), sondern seitlich. Im IPA findet sich bei den lateralen Frikativen lediglich die stimmlose Variante [ɬ] sowie das stimmhafte [ɮ].

Die beiden folgenden Artikulationsmodi gehören anders als Taps, Flaps und laterale Frikative nicht zu den pulmonalen Konoiden, da sie sich anderer Luftstrommechanismen bedienen. Beide Modi kommen in keiner europäischen Sprache als Sprachlaute vor, werden jedoch in paralinguistischer Form situativ artikuliert. Bei den so genannten **Clicks** oder *Schnalzlauten* handelt es sich um orale Verschlusslaute, die sich dem velaren Luftstrommechanismus bedienen. Clicks findet man z.B. in den südafrikanischen Khoisan-Sprachen wie dem !Xóõ. Die Artikulationsstellen der Clicks sind: bilabial [ɔ], dental [ʘ], (post)alveolar [ɕ], palatoalveolar [ɕ] und alveolar lateral [ɬ] (vgl.

¹¹So z.B. im Spanischen (vgl. Pullum & Ladusaw ²1996:161).

¹²So z.B. in Kreish (Gbaya) und Hindi (vgl. Pullum & Ladusaw ²1996:163).

3. Grundlagen der Phonetik und Akustik

Ladefoged ⁵2006:16f, Pompino-Marschall ²2003:207ff, Pullum & Ladusaw ²1996:132, 219, 221, 223, 225). Die **Ejektive** kommen insbesondere in afro-asiatischen und Indianersprachen, wie z.B. im Lakhote, einer Sioux-Sprache, vor und bedienen sich – wie bereits oben erwähnt – ebenfalls nicht des pulmonalen, sondern, anders als die Clicks, dem glottalen Luftstrommechanismus (vgl. Pompino-Marschall ²2003:209). Ejektive werden mit dem Diakritikum ['] transkribiert, also z.B. [p'] (bilabial), [t'] (dental / alveolar), [k'] (velar) oder [s'] (alveolarer Frikativ) (vgl. Clark & Yallop ²1995:17f, siehe auch Ladefoged ⁵2006:135). **Implosive** treten u.a. im Sindhi auf, einer indo-arischen Sprache die in Pakistan und Indien gesprochen wird (vgl. Pompino-Marschall ²2003:212). Sie können bilabial [ɓ], alveolar [ɗ], retroflex [ɗ̠], palatal [ɟ], velar [ɠ] oder uvular [ʁ] gebildet werden.

3.2. Akustische Phonetik und psychoakustische Grundlagen

Die physikalische Disziplin der *Akustik* ist ein Teilbereich der Mechanik bzw. der Schwingungs- und Wellenlehre und beschreibt den Ablauf und die Ausbreitung von Schwingungen in unterschiedlichen Umgebungen, wie z.B. in Gasen oder Flüssigkeiten (vgl. Kuttruff 2004:4). Sie beschäftigt sich u.a. mit den Fragen zur Entstehung des Sprechschalls durch einen Sprecher und dessen Rezeption durch einem Hörer. Um die späteren Ausführungen zur Hörwahrnehmung (vgl. Kapitel 6) und zur Individualität von Sprechstimmen (vgl. Kapitel 4.5) verstehen zu können, sollen an dieser Stelle einige grundlegende Begriffe der Akustik und der Psychoakustik¹³ erläutert werden. Dazu gehört streng genommen auch der Bereich der Sprechschallakustik mit ihren intra- und interindividuellen Merkmalen. Da sie für den weiteren Verlauf der Arbeit jedoch von besonderer Relevanz sind, sollen sie gesondert in Kapitel 4.2 und 4.3 erörtert werden.

3.2.1. Akustische Grundlagen der Schallproduktion

Im Hinblick auf die akustischen Phänomene, die während einer Schallproduktion auftreten, soll an dieser Stelle insbesondere auf die für den weiteren Verlauf der Ausführungen relevanten Begriffe *Frequenz* und *Amplitude* eingegangen werden. In einem zweiten Abschnitt werden unterschiedliche Schallformen kurz vorgestellt.

¹³Die Psychoakustik beschäftigt sich mit der systematischen Erforschung der Hörwahrnehmung (vgl. Kuttruff 2004:5).

Frequenz und Amplitude

Ein zentraler Begriff der Akustik ist *Schall*. Darunter versteht die Physik im Allgemeinen Moleküle, die „Wechselbewegungen um eine potentiellen Mittel- oder Ruhelage“ (Pétursson & Neppert ³2002:125) (potenzielle Gleichgewichtslage) vollführen und dabei minimale Druckschwankungen (Schallschwingungen) verursachen. Wie bereits in Kapitel 3.1.2 erläutert, werden diese (Luft-)Druckschwankungen durch den auf Grund des Bernoulli-Effektes immer wiederkehrenden Glottisverschluss und die daraus resultierende Kavitätenbildung verursacht. Um vom menschlichen Hörorgan wahrgenommen werden zu können, müssen die molekularen Schwingungen in einer bestimmten Geschwindigkeit ablaufen, nämlich etwa 16 bis 20.000 mal pro Sekunde, also mit einer **Periodenfrequenz**¹⁴ [f] von 16 Hz bis 20 kHz¹⁵ (vgl. Pompino-Marschall ²2003:87, Clark & Yallop ²1995:207). Dies entspricht dem *hörbaren Frequenzband* des Menschen. Die Frequenz beschreibt somit das Verhältnis von Schwingungszahl zu Zeit ($1/s = 1 \text{ Hz}$)¹⁶. Allgemeiner formuliert gilt:

$$f = \frac{1}{T} \quad (3.2)$$

Da die Periodenfrequenz [f] und die Periodendauer [T] im reziproken Verhältnis zueinander stehen, gilt auch

$$T = \frac{1}{f}. \quad (3.3)$$

Die Schallereignisse müssen jedoch nicht nur in einer bestimmten Frequenz ablaufen sondern auch eine bestimmte *Intensität* (Mindestdruck) überschreiten. Dies führt zu einem weiteren zentralen Begriff der Akustik, der **Amplitude** (A) (z.B. μm), die die positive und negative Maximalauslenkung der Moleküle von ihrer potenziellen Gleichgewichtslage beschreibt (siehe oben).

Im Gegensatz zu den lediglich hin und her schwingenden molekularen Luftpartikeln breitet sich die eigentliche Druckwelle mit einer Geschwindigkeit von etwa 340 m/s ¹⁷ als Stoß- oder Longitudinalwelle kugelförmig im Schallfeld aus (bei normalem

¹⁴Verkürzt oftmals auch lediglich als *Frequenz* bezeichnet.

¹⁵Doch auch die Frequenzen ober- und unterhalb des für den Menschen hörbaren Frequenzbandes werden zur Kommunikation genutzt. Finnwale, Elefanten und Krokodile beispielsweise nutzen Frequenzen im Bereich des Infraschalls, Insekten, Delfine und Spitzmäuse solche im Ultraschallbereich (vgl. Langenmayr 1997:34-44).

¹⁶Benannt nach dem Physiker Heinrich Rudolf HERTZ (1857-1894). In der angloamerikanischen Literatur findet sich gelegentlich auch die Abkürzung *cps* (*cycles per second*) (siehe z.B. Laver 1994:450, Baldwin & French 1990:45)

¹⁷Schallgeschwindigkeit in Luft [c_L].

3. Grundlagen der Phonetik und Akustik

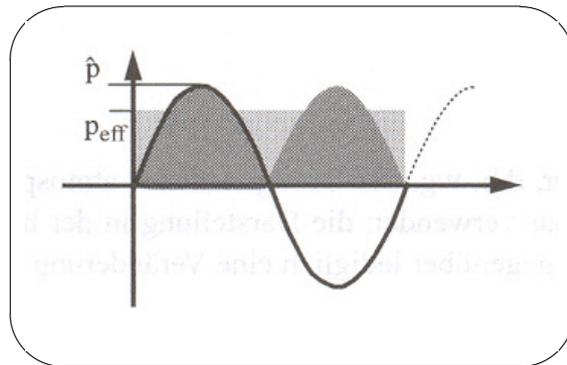


Abbildung 3.9.: Spitzen- und effektiver Schalldruck
(aus: Pompino-Marschall ²2003:88)

atmosphärischem Druck und einer Temperatur von 20°C). Die Wellenlänge $[\lambda]$ stellt somit eine frequenzabhängige „räumliche Abfolge von Druckminima und -maxima“ (Pompino-Marschall ²2003:91) nach der Formel

$$\text{Wellenlaenge} = \frac{\text{Schallgeschwindigkeit (Luft)}}{\text{Periodenfrequenz}}$$

oder

$$\lambda = \frac{c_L}{f} \quad (3.4)$$

dar (siehe oben). Dabei nimmt der effektive Schalldruck $[p_{eff}]$ (vgl. Abbildung 3.9) mit der Entfernung von der (punktuellen) Schallquelle quadratisch ab (vgl. Pompino-Marschall ²2003:91). Das bedeutet beispielsweise, dass in doppelter Entfernung vom Schallerzeuger die Intensität nur noch ein Viertel so groß ist (vgl. auch Kapitel 3.2.2).

Unterschiedliche Schallformen

Die Schallschwingungen wiederum lassen sich (psycho)akustisch in drei unterschiedliche Kategorien einteilen, die hier jedoch nur in aller Kürze erläutert werden sollen. Es handelt sich dabei um

- Töne,
- Klänge und
- Geräusche.

3. Grundlagen der Phonetik und Akustik

Bei **Tönen** führen die Moleküle eine Pendel- bzw. Sinoidalschwingung durch. Von ihnen lassen sich alle anderen Schwingungsformen ableiten. Abbildung 3.10 zeigt den Zeitverlauf einer solchen Sinoidalschwingung. Sie wird auch als *reiner Ton* bezeichnet (vgl. Kuttruff 2004:203).

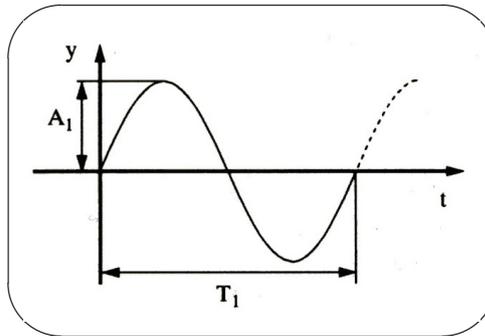


Abbildung 3.10.: Amplitudenveränderung über die Zeit
(aus: Pompino-Marschall ²2003:94)

Isoliert lautierete Vokale entsprechen psychoakustisch **Klängen** oder *komplexen Tönen* (vgl. Kuttruff 2004:221ff). Sie bestehen aus Schwingungen, die aus einzelnen Sinoidalschwingungen zusammengesetzt sind. Stehen die Sinoidalschwingungen in einem harmonischen Verhältnis zueinander, stellen sie also ein Vielfaches einer gemeinsamen Grundfrequenz (F_0)¹⁸ dar, so spricht man von Klängen im eigentlichen Sinn. Stehen die einzelnen Sinoidalschwingungen jedoch in einem beliebigen Frequenzverhältnis zueinander, wie z.B. beim Glockenklang, so spricht man von Klängen im weiteren Sinn. Zeitlich betrachtet handelt es sich bei beiden Schalltypen um periodische Schwingungen, da sie sich aus einzelnen Sinoidalschwingungen zusammensetzen. Klänge bestehen also stets aus einem Grundton und mehreren Ober-, Partial- bzw. Teiltönen. Die Gesamtheit aller Obertöne bilden das Frequenzspektrum des Tones. Abbildung 3.11 auf der nächsten Seite zeigt die additive Überlagerung (*Fourier-Synthese*) mehrerer harmonischer Sinoidalschwingungen (vgl. Rose 2002:200ff).

Handelt es sich bei den Schwingungen um zufällig verteilte Amplitudenwerte, so spricht man von einem **Geräusch** (z.B. Gewehrschuss, Wasserrauschen etc.) (vgl. Kuttruff 2005:202, Pompino-Marschall ²2003:91). Geräusche sind stets *aperiodisch*, da ihr Entstehen das Ergebnis von Zufallsprozessen ist. Dies ist in Abbildung 3.12 auf der nächsten Seite anhand eines Frikativrauschens dargestellt.

¹⁸Unter Grundfrequenz versteht man die tiefste bzw. unterste Frequenz einer komplexen Wellenform.

3. Grundlagen der Phonetik und Akustik

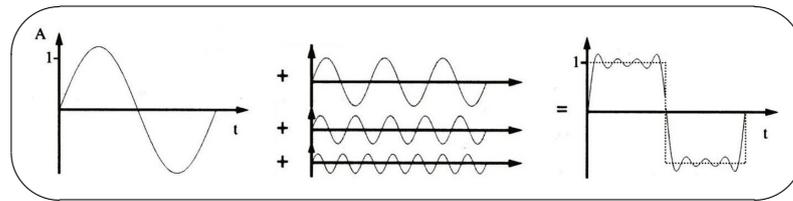


Abbildung 3.11.: Fourier-Synthese harmonischer Sinusschwingungen
(verändert nach Pompino-Marschall ²2003:95)

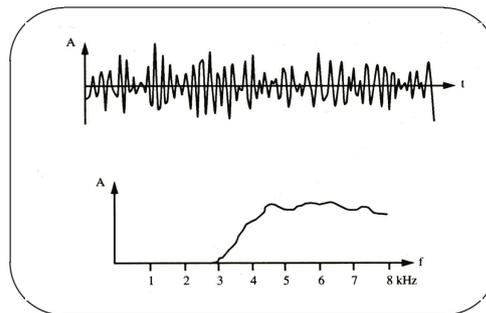


Abbildung 3.12.: Geräusch als Zeitsignal und als kontinuierliches Spektrum
(aus: Pompino-Marschall ²2003:98)

3.2.2. Schallrezeption: Wahrnehmung von Tonhöhe und Lautstärke

Schwerpunkt der Psychophysik bzw. der Psychoakustik ist die Beschreibung der Zusammenhänge zwischen den rein physikalischen Größen auf der einen und den wahrgenommenen Ereignissen auf der anderen Seite. Es geht also um die „Formulierung des mathematischen *Zusammenhangs* zwischen den *physikalischen Skalierungsgrößen* der Parameter des akustischen Signals und den zugeordneten *psychologischen Skalierungsgrößen* der auditiven Wahrnehmung“ (Pompino-Marschall ²2003:153, Hervorhebungen im Original). In den folgenden Unterkapiteln sollen die für die auditorische Sprachwahrnehmung relevanten Grundlagen dargestellt werden. Dazu gehört insbesondere die Wahrnehmung der Lautstärke und die der Tonhöhe.

Wahrnehmung der Lautstärke

Wie bereits erläutert ist die physikalische Skalierungsgröße des Schalldruckpegels [dB] mit unterschiedlichen psychoakustischen Wahrnehmungsgrößen verbunden. Die Psychoakustik unterscheidet dabei zwischen der

„frequenzabhängigen Skalierung der als gleich laut wahrgenommenen **Lautstärke** (...) und der daran anschließenden Skalierung nach dem lautstärkemäßigen Verhältnis akustischer Signale untereinander (also z.B. dop-

3. Grundlagen der Phonetik und Akustik

pelt oder halb so laut), der sog. **Lautheit**“ (Pompino-Marschall ²2003:153, Hervorhebungen im Original).

Um die Empfindung der **Lautstärke** eines Schallereignisses T_1 zu quantifizieren bittet man Versuchspersonen, den Lautstärkepegel eines 1-kHz-Sinustones so einzustellen, bis dieser die Lautstärke des akustischen Signals T_1 erreicht hat. Die Lautstärke von T_1 ist nun der Pegel des 1-kHz-Tones und wird nicht in dB sondern in *phon* angegeben. Dies bedeutet, dass die „Lautstärke eines akustischen Signals gleich der bei einem 1-kHz-Ton mit einem Schalldruckpegel gleicher Höhe [in dB] ist“ (Pompino-Marschall ²2003:153). Per definitionem stimmt bei einer Frequenz von 1 kHz der Lautstärkepegel [phon] mit dem Schalldruckpegel [dB] überein (vgl. Kutt-ruff 2004:238). Ein 1000-Hz-Sinuston mit einem Schalldruckpegel von 40 dB hat also eine Lautstärke von 40 phon. Streng genommen gilt dies heute jedoch nicht mehr sondern bezieht sich auf die ersten Messungen in den frühen dreißiger Jahren. Heutigen Messungen zufolge liegt die Kurve um ≈ 4 dB über den ursprünglich ermittelten Werten (vgl. Neppert ⁴1999:82).

Wiederholt man das oben geschilderte Verfahren mit Sinustönen anderer Frequenzen, so erhält man die so genannten *Kurven / Isophone gleicher Lautstärke*, die in Abbildung 3.13 dargestellt sind. Dabei sind die ursprünglich ermittelten Werte von Fletcher & Munson (1933) durch den blauen Linienverlauf (ISO 226), die aktuellen Werte durch den roten Linienverlauf gekennzeichnet (ISO 226:2003).

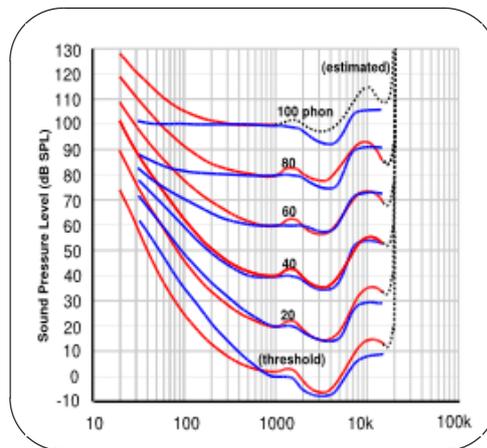


Abbildung 3.13.: Kurven gleicher Lautstärkepegel nach ISO 226 / ISO 226:2003 (mit freundlicher Genehmigung von Peter J. Skirrow, Lindos Electronics, Melton / Suffolk, United Kingdom)

3. Grundlagen der Phonetik und Akustik

Da die dB-Schritte in der oben erläuterten Phonskala den unterschiedlichen Stufen der **Lautheit** nur entfernt entsprechen, wurde eine Verhältnisskala eingeführt, die das Lautstärkeverhältnis zweier Töne ausdrücken kann (vgl. Neppert ⁴1999:85, Pompino-Marschall ²2004:154). Mit ihrer Hilfe sind also Aussagen möglich wie „Schallereignis T_1 ist doppelt so laut wie Schallereignis T_2 “ oder „Ton T_3 ist halb so laut wie Ton T_4 “ etc. Die Maßeinheit zur Bewertung und Messung der Lautheit ist *son*. Ein *son* ist definiert als die subjektiv empfundene Lautheit eines 1-kHz-Sinustones mit einer Lautstärke von 40 phon (in diesem Fall also mit einem Schalldruckpegel von ≈ 40 dB, siehe oben). Eine Erhöhung der Lautstärke um 10 phon entspricht einer Verdoppelung der Lautheit in *son*. Die Werte wurden ermittelt, in dem Probandenangaben, wann sie einen bestimmten Schall 2^n (also doppelt, vier-, acht-) mal so laut, bzw. halb, ein viertel, ein achtel so leise empfanden wie einen Referenzton. Abbildung 3.14 zeigt den psychoakustischen Zusammenhang zwischen dem Lautstärkepegel [phon] auf der einen und der Lautheit [sone] auf der anderen Seite.

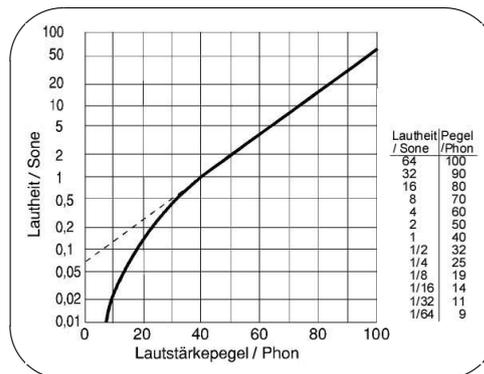


Abbildung 3.14.: Zusammenhang zwischen Lautstärkepegel und Lautheit nach ISO 532B

Wahrnehmung der Tonhöhe

Das psychoakustische Korrelat der Frequenz ist die **Tonhöhe**, die in der Einheit *mel* angegeben wird. Um herauszufinden, welche Frequenz welcher Empfindungsgröße entspricht, nutzt man ein ähnliches Verfahren wie bereits im vorhergehenden Abschnitt erläutert. Diesmal werden Probanden gebeten, zu einem vorgegebenen Ton T_1 einen weiteren Ton T_2 zuzuordnen, der als doppelt so hoch empfunden wird (vgl. Kuttruff 2004:231, Neppert ⁴1999:59). In älteren Untersuchungen (z.B. Zwicker & Feldtkeller ²1967:78ff) diente das kleine *c* (131 Hz) als Bezugspunkt, heute die auch

3. Grundlagen der Phonetik und Akustik

in der Audiometrie übliche Testfrequenz von 125 Hz. Die daraus resultierende Skala ist in Abbildung 3.15 wiedergegeben. Aus ihr wird ersichtlich, dass bis zu einer Frequenz von 0,5 kHz die mel-Werte mit den (k)Hz-Werten übereinstimmen. Dies ist bei einer weiteren Verdoppelung der empfundenen Tonhöhe auf 1000 mel nicht mehr der Fall (1000 mel = 1140 Hz, 2000 mel = 5020 Hz usw.).

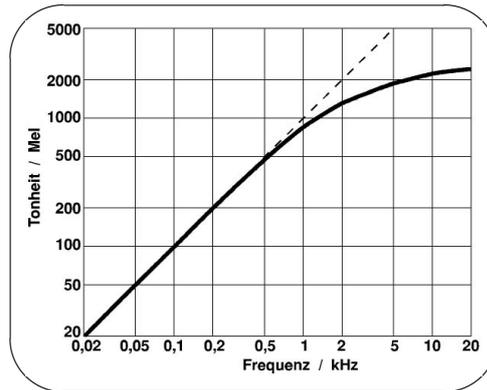


Abbildung 3.15.: Zusammenhang zwischen subjektiver Tonhöhe und Frequenz

Für die Lautheit komplexer Schälle, wie sie beispielsweise bei gesprochener Sprache vorliegen, ist die so genannte **Tonheit** von Interesse. Sie wird mit Hilfe der Einheit *Bark*¹⁹ ausgedrückt (1 Bark \approx 100 mel). Die Bark-Skala (vgl. Abbildung 3.16) wurde wie die bereits erwähnte mel-Skala mit Hilfe des kleinen *c* (131 Hz) normiert²⁰ und reicht von 0,2 bis 25 Bark.

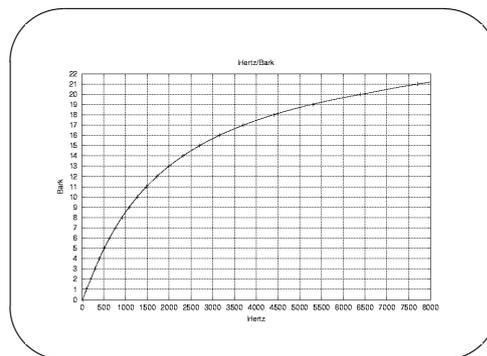


Abbildung 3.16.: Zusammenhang zwischen Bark und Frequenz
(Internetquelle: <http://www-users.york.ac.uk/~pgc104/images/barkcurve.gif> (29.12.2005))

¹⁹ Benannt nach dem Physiker Heinrich Georg BARKHAUSEN (1881-1956).

²⁰ Es gilt also 1,31 Bark \approx 131 mel \approx 131 Hz.

3. Grundlagen der Phonetik und Akustik

In der Literatur zur Akustik spielt die Einheit keine große Rolle und ist nur deshalb relevant, da auf ihr die Unterteilung des für den Menschen hörbaren Frequenzbandes in so genannte *Frequenzgruppen* beruht. Innerhalb dieser Gruppen verhalten sich die einzelnen Frequenzanteile nicht additiv sondern der stärkste Anteil (z.B. eine einzelne Teilschwingung) verdeckt weitestgehend alle anderen Anteile (vgl. Pompino-Marschall ²2004: 156, Neppert ⁴1999:62). Diese Phänomen wird als *Verdeckung* oder *Shadowing* bezeichnet. Frequenzkomponenten unterschiedlicher Frequenzgruppen werden vom Hörenden zu einer Gesamtlautheit gemittelt, so dass

„der Mittelwert aus den jeweiligen frequenzgruppenbezogenen 'spezifischen Lautheiten' [in sone / Bark] für die wahrgenommene Gesamtlautheit des Schallereignisses maßgebend (ist)“ (Pompino-Marschall ²2004:156f).

Im vorliegenden Kapitel wurden neben den für das deutsche Sprachsystem relevanten Lauten auch die anatomischen und physiologischen Grundlagen der Funktionskreise *Atmung*, *Phonation* und *Artikulation* vorgestellt. Der zweite Teil beschäftigte sich mit dem zentralen Begriff der Akustik, dem *Schall*, und den akustischen Besonderheiten des Sprechschalls. Für den weiteren Verlauf der Arbeit ist die Frage, welche konkreten akustischen Parameter nun für das Erkennen einer menschlichen Stimme genutzt werden, von zentraler Bedeutung. Daher steht die bereits weiter oben erläuterte *Frequenz* eines Schallsignals im Mittelpunkt des folgenden Kapitels, da sie die maßgeblichen Informationen für das Identifizieren einer Stimme, insbesondere im Hinblick auf die *mittlere Sprechstimmlage* (vgl. Kapitel 4.3.1) und die *Intonation* (vgl. Kapitel 4.3.2), übermitteln kann.

4. Suprasegmentalia in der Spontansprache

4. Suprasegmentalia in der Spontansprache

„The voice is the very emblem of the speaker, indelibly woven into the fabric of speech. In this sense, each of our utterances of spoken language carries not only its own message, but through accent, tone of voice and habitual voice quality it is at the same time an audible declaration of our membership of particular social regional groups, of our individual physical and psychological identity, and of our momentary mood“ (Laver 1994:2).

Während im vorangegangenen Kapitel der Fokus auf isolierten Lauten lag, richten wir unseren Blick nun auf die zu einem akustischen Kontinuum verketteten Einzellaute in größeren Laut- bzw. Segmentgruppen. Dabei fällt auf, dass bei einer spontansprachlichen Äußerung nicht nur Lautabfolgen produziert werden, sondern darüber hinaus zahlreiche weitere lautliche Erscheinungen auftreten. Dazu zählen z.B. Rhythmus, Intonation, Akzentuierung, Lautstärke, Stimmqualität etc. Diese lautlichen Phänomene, die sich nur im Rahmen einer Sequenz bestimmen lassen, sind relative Größen und an längere Äußerungseinheiten – wie Silben, Wörter oder Sätze – geknüpft. Da sie sich also über mehr als nur ein Segment erstrecken, werden sie unter dem Terminus *Suprasegmentalia* zusammengefasst (vgl. Willi 2001:422).

Manche Suprasegmentalia übernehmen dabei auch sprachliche Funktionen. So beispielsweise die **Intonation**. Sie basiert auf spezifischen Grundfrequenzverläufen und kann grammatikalische Kategorien ausdrücken. Im Deutschen realisiert dies ein Sprecher dadurch, indem er die Satzmelodie variiert. Durch die Veränderung des satzfinalen Tonhöhenverlaufs kann so beispielsweise aus einer Aussage eine Frage entstehen:

Das Kind spielt mit dem Bagger.

vs.

Das Kind spielt mit dem Bagger?

Auch die **Akzentuierung** einer Silbe innerhalb eines Wortes kann eine sprachliche Funktion erfüllen.

erlangen vs. *Erlangen*

umfahren vs. *umfahren* (vgl. Willi 2001:423)

Konstanz vs. *Konstanz* (vgl. Alter 2002:150).

Für die vorliegende Arbeit sind insbesondere jene Suprasegmentalia von Bedeutung, die Hinweise auf Eigenschaften der Sprechenden Person geben, also Informationen über Alter, Geschlecht, psychische Verfassung etc. übermitteln. Anhand solcher lautlicher Erscheinungen gelingt es uns auch, einer gehörten Stimme einen spezifischen Sprecher zuzuordnen. Welche Suprasegmentalia und akustischen Korrelate ein

4. *Suprasegmentalia in der Spontansprache*

Hörer dabei nutzt, wird im folgenden Kapitel ausführlich analysiert und dargestellt. Die Störungen der prosodischen Verarbeitung, die zum Formenkreis der auditiven Agnosien gezählt werden, sind Inhalt des Kapitels 7.3.

4.1. **Erkennungsrelevante Sprechermerkmale**

Einer der wohl ältesten Berichte, der sich mit „voice identification“ und „aural-perceptual testimony“ (Hollien 1990:192) beschäftigt, stammt von einem englischen Gericht aus dem Jahre 1660. Als historisches Beispiel für die Geburtsstunde der (zunächst forensisch motivierten) Sprechererkennung wird die Entführung und anschließende Ermordung des Sohnes der US-amerikanischen Fliegerlegende Charles Lindbergh im Jahre 1932 angeführt. Für diese Tat wurde der deutschstämmige Bruno Richard Hauptmann zum Tode verurteilt und im März 1936 hingerichtet. Als besonders schwerwiegenden Beweis werteten die Richter seinerzeit die Aussage des Oberst Lindbergh, der drei Jahre nach der Tat in der Stimme Hauptmanns die des Anonymus erkannt haben will, der die telefonischen Lösegeldforderungen an ihn gerichtet hatte (vgl. McGehee 1937:249, Künzel 1988:216f, 1990:1, 1997:101, Hollien 1990:195). Diese Behauptung löste in kriminalwissenschaftlichen Fachkreisen eine kontroverse Diskussion über die menschliche Fähigkeit der Stimmerkennung aus. McGehee nahm sich anlässlich der Verurteilung Hauptmanns als eine der ersten der Problematik an und führte mehrere Untersuchungen zu unterschiedlichen Fragestellungen der Stimmerkennung durch. So zeigte sie unter anderem, dass lediglich 83 % der teilnehmenden Versuchspersonen einen Tag nach dem Hören einer unbekanntenen Stimme diese noch korrekt wieder erkennen konnten. Nach zwei Wochen waren dazu nur noch 51 % und nach fünf Monaten nur noch 13 % in der Lage (McGehee 1937:262, 1944:53). Als Ergebnis der Forschungsbemühungen in den folgenden Jahrzehnten insbesondere in den Bereichen der (Neuro-) Phonetik, der Neuropsychologie, der Neurolinguistik und der Forensik haben sich im Laufe der vergangenen Jahrzehnte drei für die Individualität einer sprechenden Person relevante Merkmalsbereiche heraus kristallisiert. Es handelt sich dabei um Merkmale aus den Bereichen

- **Stimme,**
- **Sprache** und
- **Sprechweise,**

auf die in den nun folgenden Unterkapiteln näher eingegangen werden (vgl. Schall 2004:548f, Künzel 1998:384, Jessen 2004).

4. *Suprasegmentalia in der Spontansprache*

An dieser Stelle sei darauf hingewiesen, dass die Beschreibung der individuellen Sprechermerkmale einer integrierender Betrachtung der bereits an anderer Stelle beschriebenen artikulatorischen (vgl. Kapitel 3.1) und akustischen Phonetik (vgl. Kapitel 3.2) bedarf. Zum besseren Verständnis werden entsprechende Sachverhalte in aller Kürze wiederholt.

4.2. Die Problematik der Inter- und Intravariabilität von Stimmen

Hören wir z.B. durch eine geöffnete Tür die Stimmen unterschiedlicher Kollegen oder Freunde aus einem angrenzenden Raum, so gelingt es uns in den meisten Fällen, den wahrgenommenen Stimmen eine Sprecheridentität zuzuweisen. Dabei ist nahe liegend, dass dies leichter fällt, je größer die Unterschiede zweier Stimmen sind (z.B. alter Mann vs. junge Frau o.ä.). Die Ursachen der **Intersprechervariabilität** (*between speaker variability*) liegen zum einen in den organischen und anatomischen Unterschieden des Larynx, des Ansatzrohres und der Stimmlippen, zum anderen in unterschiedlichen erlernten Faktoren, wie dem Gebrauch der Artikulationswerkzeuge, dem Dialekt und anderen idio- und soziolektalen Merkmalen. Während die organischen Unterschiede durch Vererbung, Geschlecht und Alter determiniert sind, werden die erlernten Faktoren von geographischen, sozialen und kulturellen Faktoren beeinflusst (vgl. Hecker 1971:4, Neppert ⁴1999:193, Nolan 1983:26, Tosi 1979:55). Die verursachenden Faktoren der Intersprechervariabilität werden in den Kapiteln 4.3 bis 4.5 ausführlich erläutert.

Gleichzeitig zeigt uns unsere Alltagserfahrung jedoch auch, dass es neben den interindividuellen Unterschieden auch eine intraindividuelle Variabilität (*within speaker variability*) geben kann. Mit der Problematik dieser **Intrasprechervariabilität** wurde man zunächst in den frühen sechziger Jahren konfrontiert, als man begann, sich mit der maschinellen Sprechererkennung (*speaker recognition by machine, SRM*) auseinanderzusetzen. So konnten Hargreaves & Starkweather bereits 1963 zeigen, dass zahlreiche Erkennungsfehler nur deshalb entstanden, weil die Referenzmuster nicht am gleichen Tag sondern bereits zu einem früheren Zeitpunkt aufgenommen worden waren (1963:66).

4. Suprasegmentalia in der Spontansprache

Das Phänomen der Intrasprechervariabilität soll in Abbildung 4.1 durch die graphisch dargestellte *mittlere Grundfrequenz* (blaue Linie), die jeweils ersten drei *Formanten* (rote Punkte) und die *Intensität* (grüne Linie) illustriert werden.

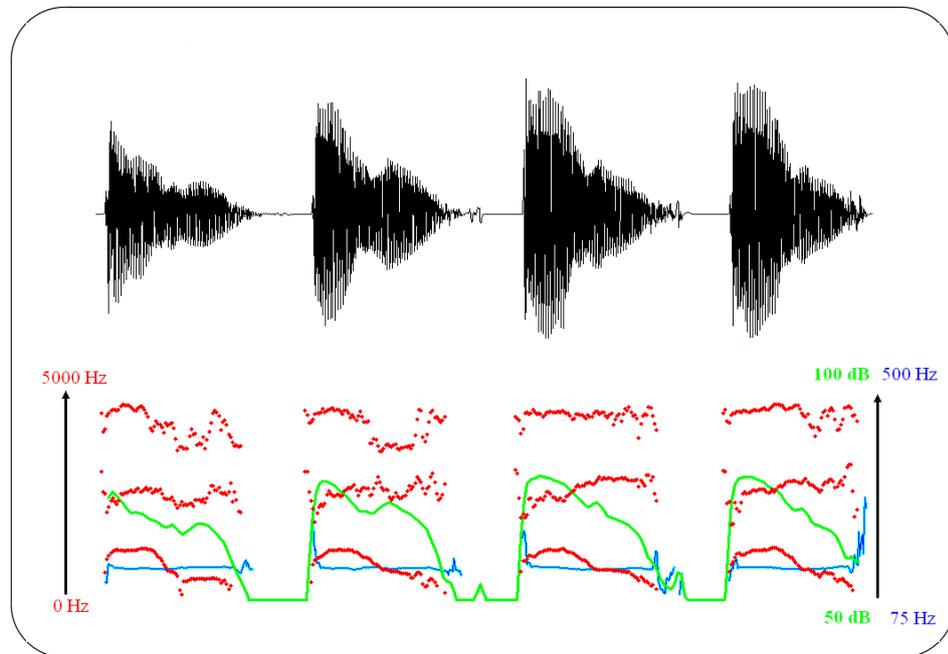


Abbildung 4.1.: Intraindividuelle Unterschiede am Beispiel <Ball> (männlicher Sprecher; eigene Aufnahme; Analyse und Darstellung mit Praat)

In diesem Beispiel artikuliert ein männlicher Sprecher im Abstand von jeweils zwei Sekunden das Wort <Ball> (die Pausen wurden aus Darstellungsgründen entfernt). Die Instruktion lautete, das Wort möglichst auf gleiche Art und Weise zu wiederholen (also ohne unterschiedliche Betonungen u.ä.). Wie deutlich sichtbar, ähneln sich die ermittelten Daten der einzelnen Äußerungen sehr stark, sie sind jedoch zu keinem Zeitpunkt genau identisch¹. In Tabelle 4.1 auf der nächsten Seite sind die wichtigsten Messergebnisse der Stimmanalysen gegenübergestellt.

Ein Aspekt der Intrasprechervariabilität sind die so genannten **Mikrovariationen**. Zwar sind sie für den eigentlichen Kommunikationsprozess irrelevant und auditiv nur schwer bestimmbar, dennoch eignen sie sich, um das Phänomen der Intrasprechervariabilität zu verdeutlichen. So konnten Orlikoff & Baken (1989a, 1989b) bei gesunden Probanden im Rahmen von Tonhalteübungen gleichmäßige Schwankun-

¹Weitere Beispiele dazu finden sich bei Hollien (1990:216-219) und Rose (2002:232-237), siehe auch Atkinson (1976).

4. Suprasegmentalia in der Spontansprache

| | Version | | | |
|---------------------------|---------|---------|---------|---------|
| | 1 | 2 | 3 | 4 |
| mittl. Grundfrequenz (Hz) | 128,094 | 125,82 | 123,582 | 130,107 |
| F ₀ min. (Hz) | 116,746 | 113,492 | 88,374 | 105,475 |
| F ₀ max. (Hz) | 189,125 | 185,176 | 162,722 | 241,124 |
| stimml. Anteil (%) | 2,055 | 2,041 | 1,325 | 12,752 |
| Jitter (%) | 2,055 | 2,041 | 2,659 | 12,752 |
| Shimmer (%) | 3,525 | 4,354 | 3,665 | 3,378 |
| mittl. Intensität (dB) | 64,92 | 67,789 | 68,851 | 69,28 |

Tabelle 4.1.: Intraindividuelle Unterschiede (Messergebnisse der Stimmenanalyse)
(Sprecher wie in Abbildung 4.1 auf der vorherigen Seite)

gen der Grundfrequenz nachweisen (*Jitter*). Sie wiederholten sich periodisch, wobei die Periodendauer ungefähr dem zeitlichen Abstand zweier Pulsschläge entsprach. Dies lässt die Vermutung zu, dass mit jedem Pulsschlag die Stimmlippen in minimalem Umfang aufgebläht werden und somit ihr Schwingungsverhalten periodisch verändern (vgl. Orlikoff & Baken 1989b:581). Titze (1991) konnte zeigen, dass die Impulse aus dem ZNS ein Zusammenziehen des *Musculus thyroarytaenoideus* und anderer laryngealer Muskeln bewirken, was ebenfalls Einfluss auf das Schwingungsverhalten der Stimmlippen habe und somit einen physiologischen *Jitter*² erklären würde (vgl. Titze 1991:470). Linville & Korabic (1987) und Linville (1988) wiesen in ihren Untersuchungen nach, dass auch das Alter Einfluss auf das Vorhandensein von Mikrovariationen haben kann. Beim Vergleich von älteren mit jüngeren Sprecherinnen zeigte sich, dass beispielsweise ältere Frauen (69-90 Jahre) beim Lautieren des Vokoids [a:] einen höheren *Jitter* aufwiesen als bei den Vokoiden [i:] und [u:]. Bei jüngeren Frauen (18-22 Jahre) zeigten sich exakt umgekehrte Verhältnisse (vgl. Linville 1988:743f, Linville & Korabic 1987:1197f).

Wie bereits an anderer Stelle erläutert, sind die Stimmbildungs- und Artikulationsprozesse sehr komplex. Die oben erläuterten Mikrovariationen und kleinere Abweichungen in der Atmung bzw. der laryngealen und artikulatorischen Aktivität spielen für den Erfolg eines Kommunikationsprozesses in der Regel keine Rolle, so dass Sprecher und Hörer nicht auf kleine 'Unregelmäßigkeiten' achten. Diese physiologischen Schwankungen führen jedoch dazu, dass die individuellen akustischen Signale keine

²Schwankungen der Grundfrequenz.

4. Suprasegmentalia in der Spontansprache

unveränderliche Größe sind, sondern in einem in der Regel eng gesteckten Toleranzbereich liegen.

4.3. Interindividuelle Merkmale der Stimme

Der Merkmalsbereich **Stimme** liefert – insbesondere für die Fragestellungen der Untersuchungen die im zweiten Teil dieser Arbeit beschrieben werden sollen – die wichtigsten Faktoren bezüglich der Individualität einer Stimme. Als relevante Faktoren sind hier insbesondere

- die mittlere Sprechstimmlage,
- der Intonationsverlauf und
- die Stimmqualität

zu nennen (vgl. Künzel 1998:385, 1997:104, Jessen 2004).

4.3.1. Die mittlere Sprechstimmlage

Eine wichtige, interindividuell sehr bedeutsame sprecherspezifische Information liefert die *mittlere Sprechstimmlage*³ (vgl. Rose 2002:244, Baldwin & French 1990:48, Hollien 1990:196, Van Lancker, Kreiman & Emmorey 1985:19, Pollack, Pickett & Sumbly 1954:404). Mit ihr wird ein relativ eng abgrenzbarer Frequenzbereich bezeichnet in dem sich die Äußerungen eines Sprechers im Normalfall bewegen und der auf Grund seiner Determiniertheit insbesondere durch die Anatomie

- der Stimmlippen und
- des Ansatzrohres

nur eine geringe intraindividuelle Variabilität aufweist (vgl. z.B. Atkinson 1976:445). Die Grundfrequenz einer Stimme wird durch die im Folgenden vorgestellten Faktoren beeinflusst.

Einfluss von Stimmlippenlänge und -masse auf F_0

Bereits bei der Erzeugung des Rohschalls kommt es zu entscheidenden interindividuellen Unterschieden. Wie bereits an anderer Stelle erwähnt, entsteht durch das Ausströmen der pulmonal-exhalatorischen Luft an der Glottis der Stimmtönen, der

³Auch *mittlere Grundfrequenz* oder *mean pitch* genannt.

4. Suprasegmentalia in der Spontansprache

im weiteren Verlauf z.B. zu einem Sprachlaut modifiziert werden kann (vgl. Kapitel 3.1). Die Frequenz dieses Stimmtones (Grundfrequenz $[F_0]$), wird durch die Rate der Stimmlippenschwingungen pro Sekunde bestimmt (vgl. Kapitel 3.2). Dabei wirkt sich neben der Länge und Elastizität auch die wirksame Masse der Stimmlippen auf deren Schwingungsverhalten aus und beeinflusst die Grundfrequenz und somit auch die *Tonhöhe*, die das vom Hörer wahrgenommene psychoakustische Korrelat zu F_0 darstellt (vgl. Kapitel 3.2.2). Das Verhältnis zwischen F_0 und der *Stimmlippenlänge* definiert Titze (1994:200, 1995:38) wie folgt:

$$F_0 = \frac{1}{2L_m} \times \sqrt{\frac{\sigma_c}{\rho}} \quad (4.1)$$

Dabei gilt L_m als Länge der Stimmlippen (in Meter), σ_c als ihre aktuelle Spannung (in Pascal) und ρ als die Gewebedichte (1040 kg / m^3). Daraus wird deutlich, dass F_0 mit der Länge der Stimmlippen abnimmt. Bei einer angenommenen Gewebedichte von 1040 kg / m^3 und einer Spannung von 15000 Pa wird eine Stimmlippe von 15 mm Länge in einer Grundfrequenz von etwa 127 Hz schwingen⁴, während eine um 3 mm längere Stimmlippe (18 mm) nur noch eine Grundfrequenz von 105 Hz produzieren würde⁵. Formel 4.1 hat jedoch nur dann Gültigkeit, wenn sie auf zwei unterschiedliche Sprecher bezogen wird, also die interindividuellen Unterschiede beschreiben soll. Für die intraindividuelle Beeinflussung von F_0 , für die der Sprecher seine Stimmlippen aktiv strecken muss, gilt der genau umgekehrte Fall: längere Stimmlippen bewirken aufgrund der aktiven Kontraktion eine höhere F_0 (vgl. Rose 2002:244).

Tabelle 4.2 gibt einen Überblick über die durchschnittliche Stimmlippenlänge und Grundfrequenz bei Frauen, Männern und Säuglingen.

| | mittl. Stimmlippenlänge | Grundfrequenzmittelwert | Streuung |
|----------|-------------------------|-------------------------|--------------|
| Frau | 13 - 17 mm | 200 - 230 Hz | 180 - 450 Hz |
| Mann | 17 - 24 mm | 110 - 130 Hz | 90 - 220 Hz |
| Säugling | 5 mm | 400 Hz | 300 - 700 Hz |

Tabelle 4.2.: Übersicht Grundfrequenzen

(vgl. Pétursson & Neppert ³2002:137, Pompino-Marschall 1995:35, Titze 1994:170)

⁴ $F_0 = \frac{1}{2 \times 0,015} \times \sqrt{\frac{15000}{1040}} = 126,592 \approx 127 \text{ Hz}$.

⁵ $F_0 = \frac{1}{2 \times 0,018} \times \sqrt{\frac{15000}{1040}} = 105,493 \approx 105 \text{ Hz}$.

4. Suprasegmentalia in der Spontansprache

Für den interindividuellen Einfluss der *Stimmlippenmasse* auf F_0 formuliert Titze (1994:193):

$$F_0 = \frac{1}{2\pi} \times \sqrt{\frac{k}{m}} \quad (4.2)$$

In Formel 4.2 gilt m als Masse (in Kilogramm) der schwingenden Stimmlippen und k als ihre Steifigkeit (in Newton / Meter). Auch hier wird deutlich, dass sich eine größere Masse in einer niedrigeren Grundfrequenz niederschlägt.

Zusammenfassend lässt sich festhalten, dass die in unterschiedlicher Tonhöhe wahrgenommene Grundfrequenz verschiedener Sprecher wesentlich von Stimmlippenlänge und -masse abhängig ist. Diese Parameter nutzt eine hörende Person, um einer Stimme beispielsweise Informationen über das Geschlecht des Sprechers zu entnehmen (vgl. u.a. Perry, Ohde & Ashmead 2001:2996, Whiteside 1998:583, Günzburger 1984:48, Lass et al. 1979:677 und Weinberg & Bennett 1971 (bei Kindern)). Trotz der geschilderten unvermeidbaren anatomischen Vorgaben ist ein Sprecher bewusst und unbewusst in der Lage, seine Grundfrequenz in einem gewissen Umfang aktiv zu beeinflussen. Dies ist z.B. möglich indem er die Stärke des Ausatemdruckes oder die Muskelspannung innerhalb seiner Stimmlippen variiert (vgl. Kuttruff 2004:219, 221, Rose 2002:216, Titze 1995:42ff, 1989:904).

Einfluss des supralaryngealen Vokaltraktes auf F_0

Wie bereits erwähnt, passiert das akustische Signal nach seiner Erzeugung die unterschiedlichen supralaryngealen Hohlräume: den Pharynxraum, den Mundhohlraum und den Nasenhohlraum.

Diese Hohlräume – deren Gesamtheit als *Ansatzrohr* bezeichnet wird – weisen, wie alle anderen Hohlräume unserer Umgebung, mehrere Eigenfrequenzen bzw. Resonanzen auf. Das individuell 'geformte' Ansatzrohr wirkt somit als akustisches Filter, das einige Frequenzen des an der Glottis entstandenen Signals verstärkt, andere hingegen abschwächt. Da die Gestalt der relevanten Hohlräume von den individuellen anatomischen Verhältnissen des Sprechenden abhängig ist, zeigen sie nur eine geringe intraindividuelle Variabilität. Ihr Resonanzverhalten ist interindividuell jedoch sehr unterschiedlich und nimmt Einfluss auf die spezifische Charakteristik einer Sprecherstimme (vgl. Kuttruff 2004:221, Belin, Zatorre & Bédard 2004:129, Rose 2002:38, Hollien 1990:41f). Dieses physikalische Phänomen erklärt FANT (1960) mit Hilfe seiner *Quelle-Filter-Theorie*.

4. Suprasegmentalia in der Spontansprache

Abbildung 4.2 zeigt auf der linken Seite zunächst das harmonische Linienspektrum des Rohschallsignals (a) und in der Mitte die spektrale Charakteristik des Filters (b), also das Spektrum der Impulsantwort. Als Ergebnis ist das Linienspektrum des Ausgangssignals dargestellt (c). Das Sprachsignal (hier $Y(n)$) ist also das akustische Ergebnis der (individuellen) Signalquelle (Quelle, hier $X(n)$) und der individuellen Modifikation (Filter, hier $H(n)$) (vgl. auch Fant 1960:19f, Clark & Yallop ²1995:236f, ausführlich in Neppert ⁴1999:136-140). Die Quelle-Filter-Theorie nach Fant ist die einzige Theorie, die diesen Vorgang beschreibt und wurde bisher nicht falsifiziert (vgl. Rose 2002:207, Fitch 2000).

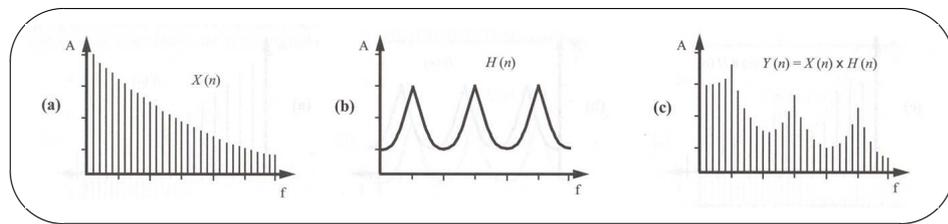


Abbildung 4.2.: Zusammenhang zwischen Quellsignal und Filtercharakteristik (aus: Pompino-Marschall ²2003:103 (verändert))

Die bisherigen Ausführungen zur Individualität einer Stimme zeigen, dass die vorgegebene Anatomie der Stimmlippen und des Ansatzrohres einen großen Einfluss auf das Klangbild einer Stimme haben und sie die auditiv wahrnehmbare Tonhöhe maßgeblich bestimmen. Diese anatomischen Verhältnisse sind nicht beliebig veränderbar und resultieren in den so genannten organisch bedingten Merkmalen, die wichtige Informationen für das Erkennen einer Stimme liefern. Dies soll jedoch nicht darüber hinwegtäuschen, dass die mittlere Grundfrequenz durch weitere ganz unterschiedliche Faktoren beeinflusst werden kann.

Weitere beeinflussende Faktoren auf die mittlere Sprechstimmlage

Rein physiologische Veränderungen der mittleren Grundfrequenz ergeben sich z.B. auf Grund des entwicklungsbedingten Größenwachstums des menschlichen Körpers (vgl. u.a. Helfrich 1979:80f). Dabei treten die größten Veränderung bei weiblichen Sprecherinnen nach dem Eintreten der ersten Menarche, also am Ende der pubertären Entwicklung ein (vgl. Pedersen et al. 1990:23, Duffy 1970:23, Hollien & Paul 1969:121, Berendes 1962:269f). Ab diesem Zeitpunkt kann man von einer adulten F_0 -Frequenz sprechen, die dann aber spätestens ab der Menopause wieder abnimmt (vgl. Amir &

4. Suprasegmentalia in der Spontansprache

Biron-Shental 2004:181f, Meurer et al. 2004:371, Schneider et al. 2004:151, Abitbol, Abitbol & Abitbol 1999:425, Boulet & Oddens 1996:17ff, Krook 1988, zu anderen Ergebnissen kommen McGlone & Hollien 1963:170). Dies bestätigt auch die Longitudinalstudie von Endres, Bambach & Flössler (1971). Bei männlichen Sprechern zeigt sich ein Erreichen der adulten Grundfrequenz im Alter von etwa 15 Jahren (Pedersen et al. 1986:26, Hollien, Green & Massey 1994:2646ff, Böhme & Hecker 1970). Im Anschluss nimmt sie bis zu einem Alter von etwa 40 Jahren kontinuierlich ab, um dann zwischen dem 60. und 80. Lebensjahr wieder anzusteigen (vgl. Gorham-Rowan & Laures-Gore 2006, Verdonck-de Leeuw & Mahieu 2004:200f, Ferrand 2002:485f, Helfrich 1979:80, Hollien & Shipp 1972:156, Shipp & Hollien 1969, Ptacek & Sander 1966, Mysak 1959:53).

Darüber hinaus ist die Grundfrequenz auch von situativen, nicht-organischen Bedingungen abhängig. Da die im Folgenden genannten Faktoren für die vorliegende Untersuchung keine Rolle spielen, sollen sie nur in aller Kürze erwähnt werden.

Zu den nicht-organischen Faktoren gehören u.a. Aufregung, Angst, lauter Störschall in der Umgebung des Sprechenden und natürlich auch Krankheit, Erkältung oder der Konsum von Alkohol oder Drogen (Rose 2002:253, 292f, Hollien 1990:191, Baldwin & French 1990:47). So beschäftigen sich beispielsweise einige Untersuchungen mit dem **Einfluss des Rauchens** auf die weibliche und männliche Stimme (vgl. u.a. Guimaraes & Abberton 2005:190, Verdonck-de Leeuw & Mahieu 2004:201, Murphy & Doyle 1987) und berichten nahezu übereinstimmend von einem Absinken der mittleren Grundfrequenz bei rauchenden Frauen bzw. Männern. So untersuchten beispielsweise Gonzalez & Carpi (2004) die Stimmen von insgesamt 134 rauchenden und nicht-rauchenden Frauen und Männern. Die mittlere Grundfrequenz der nicht-rauchenden Frauen betrug bei der genannten Untersuchung 206,4 Hz, die der rauchenden Frauen 192,4 Hz ($p < ,01$). Bei den männlichen Teilnehmern lag sie bei 125,4 Hz bzw. bei 119,4 Hz ($p < ,1$). Auch bei den anderen untersuchten Parametern (z.B. *lowest / highest fundamental frequency*, *Jitter*, *Shimmer* etc.) zeigten sich signifikante Unterschiede. Anders wirkt sich der **Konsum von Alkohol** auf z.B. die mittlere Grundfrequenz aus. So zeigen Künzel, Braun & Eysholdt (1992:44), dass die mittlere Grundfrequenz bei Männern mit einem Blutalkoholspiegel von $> 0,8 ‰$ um 2,8 Hz ansteigt ($p > ,01$) (vgl. auch Pisoni & Martin 1989, Klingholz, Penning & Liebhardt 1988:934). Übereinstimmung herrscht bei der Beobachtung, dass es durch die **Behandlung mit Hormonen** (z.B. nach der Gabe von Androgenen bei der Behandlung Transsexueller) zu Veränderungen der Grundfrequenz kommen kann. So

4. Suprasegmentalia in der Spontansprache

untersuchten u.a. Akcam et al. (2004) die stimmlichen Veränderungen bei Patienten mit einer hypogonadotrophischen Hypogonadie⁶ und den Einfluss des Androgens Testosteron. Die mittlere Grundfrequenz der Patienten lag vor der Behandlung bei 229 Hz, die der gesunden weiblichen bzw. männlichen Kontrollprobanden bei 256 Hz ($p < ,01$) bzw. 150 Hz ($p < ,001$). Nach der Behandlung mit dem Sexualhormon näherte sie sich der mittleren Grundfrequenz gesunder männlicher Kontrollpersonen an und lag bei 173 Hz ($p > ,0001$) (vgl. Akcam et al. 2004:1587, siehe auch Gooren 2005:33f, Abitbol, Abitbol & Abitbol 1999:433ff, Günzburger 1995:347, Spencer 1988:40f).

Doch auch psychologische Faktoren können stimmliche Parameter beeinflussen. So zeigen beispielsweise mehrere Untersuchungen übereinstimmend, dass, unabhängig von der Geschlechtszugehörigkeit, Patienten die unter einem **Depressionssyndrom** leiden i.d.R. eine niedrigere mittlere Grundfrequenz zeigen als nicht-depressive Kontrollpersonen (vgl. u.a. Breznitz 1992, Nilsonne 1988:253, Nilsonne 1987:235). Genau umgekehrt verhält es sich bei Patienten mit **Schizophrenie**, deren mittlere Grundfrequenz höher liegt als die der gesunden Kontrollpersonen (vgl. u.a. Covington et al. 2005:92f). Auch **Emotionen** zeigen entsprechende Einflüsse. Dabei handelt es sich entweder um simulierte Emotionen durch Schauspielerinnen und Schauspieler (siehe dazu z.B. Whiteside 1999:1195) oder auch um reale (z.B. Stress-) Situationen (siehe z.B. Protopapas & Lieberman 1997:2267). In beiden Fällen können signifikante Abweichungen von der mittleren Grundfrequenz, der Formantverteilung und anderer akustischer Parameter nachgewiesen werden (vgl. u.a. Erickson et al. 2006, Johnstone et al. 2005). Die **Ermüdung der Stimme**, z.B. bei Lehrern oder Schauspielern, wirkt sich ebenfalls auf die mittlere Grundfrequenz aus. So messen Novak et al. (1991:75) im Anschluss an eine Theateraufführung ein Ansteigen von F_0 bei männlichen um 2,7 Hz, bei weiblichen Schauspielern ein Absinken um 7,3 Hz (zum Teil widersprüchliche Beobachtungen finden sich u.a. bei Schneider et al. 2006, Jonsdottir, Laukkanen & Vilkmán 2002:282, Sodersten et al. 2002:356, Rantala, Vilkmán & Bloigu 2002:344).

⁶Unvollkommene oder fehlende Ausbildung bzw. sekundäre Rückbildung der primären und ggf. sekundären Geschlechtsmerkmale als Ausdruck einer fehlenden oder unzureichenden Funktion der Hoden bzw. Ovarie (vgl. Pschyrembel ²⁵⁹2002:756).

4.3.2. Der Intonationsverlauf

Wie bereits angedeutet, spricht ein Sprecher unter normalen Umständen (also ohne die genannten oder ähnliche stimmverändernden Einflüsse) in einer bestimmten Stimmtonfrequenz, die sich in einem vorwiegend anatomisch determinierten Frequenzbereich bewegt (mittlere Sprechstimmlage). Würde ein Sprecher seine mittlere Sprechstimmlage über mehrere Phrasen nahezu konstant beibehalten, so beschrieb ein Hörer diese Stimme vermutlich als 'monoton' bzw. 'eintönig' (im wahrsten Wortsinne). Wie sich im Alltag zeigt, kommt es im Rahmen natürlichsprachlicher Äußerungen jedoch gewöhnlich an bestimmten Stellen zu Abweichungen von der mittleren Sprechstimmlage, um z.B. Akzente zu setzen oder bestimmte Satz- oder Wortteile zu betonen⁷ (vgl. z.B. Künzel 1987:82). Dieser, *Intonation* oder *Satzmelodie* genannte Verlauf drückt sich u.a. in dem mathematischen Streuungsmaß der Grundfrequenz aus und wird daher auch als *long-term fundamental frequency* oder kurz LTF_0 bezeichnet (vgl. Rose 2002:248). Dies soll mit Hilfe von Abbildung 4.3 auf der nächsten Seite verdeutlicht werden. Sie zeigt u.a. die LTF_0 einer weiblichen Sprecherin, die den Satz <Die Lehrerin ermahnt die beiden Schüler.> artikuliert. Der durch die rot gestrichelte Linie gekennzeichnete Mittelwert von F_0 beträgt hier 179,6 Hz (siehe Pfeil „mean pitch“) mit einer Spanne von 76,3 Hz (siehe Pfeil „min. pitch“) bis 247,6 Hz (siehe Pfeil „max. pitch“).

Dabei hängt die Streuung von der absoluten Höhe des Mittelwertes ab und ist bei hohen Mittelwerten im Mittel größer als bei niedrigen Mittelwerten (hier: 37,2 Hz) (vgl. Künzel 1987:83). Aus diesem Grund empfiehlt es sich, sie als prozentualen Anteil der LTF_0 auszudrücken (hier: 20,78 %). Dieser Wert wird *Variationskoeffizient* [V_{koeff}] genannt.

Der Grundfrequenzmittelwert im Rahmen einer eigenen unveröffentlichten Untersuchung betrug bei gesunden weiblichen Sprechern ($N = 10$) 196,4 Hz, bei den männlichen Sprechern ($N = 11$) 122,6 Hz ($p < .001$). Wie bereits weiter oben erläutert zeigt die Erfahrung, dass die Standardabweichungen von LTF_0 bei höheren Mittelwerten größer ausfällt als bei niedrigeren Werten. Auch dies konnte in der Untersuchung bestätigt werden. So betrug die Standardabweichung bei den weiblichen Probanden im Mittel 36,6 Hz, bei den männlichen Probanden 22,4 Hz ($p = .007$). Setzt man die Standardabweichungen jedoch in Bezug zur LTF_0 , so ergibt dies bei befundfreien Sprecherinnen und Sprechern einen Variationskoeffizienten (siehe oben) von 18,77 % bzw. 18,43 % ($p = .876$).

⁷Ein anschauliches Beispiel zeigt Rose 1999a, 1999b und 2002:426ff anhand des Wortes <Hello> (siehe auch Langenmayr 1997:267).

4. Suprasegmentalia in der Spontansprache

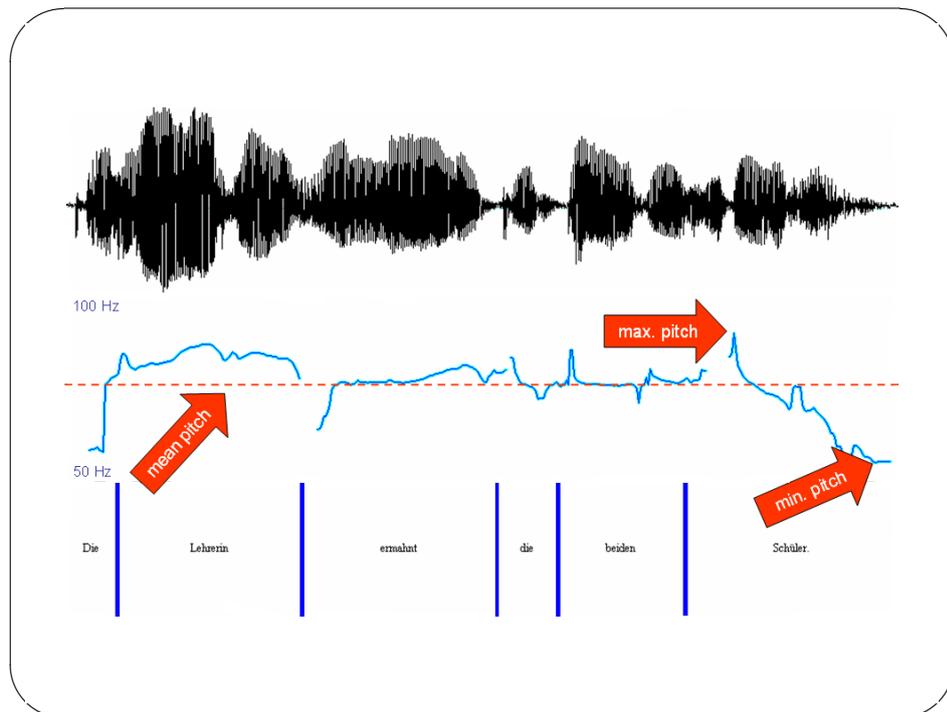


Abbildung 4.3.: Beispiel für einen Intonationsverlauf
(weibliche Sprecherin; eigene Aufnahme)

Wie die mittlere Sprechstimmlage, ist auch der Intonationsverlauf von situativen Bedingungen, wie Angst, Aufregung etc. abhängig und willentlich beeinflussbar. Da er auf dem anatomisch determinierten Grundfrequenzmittelwert beruht, zählt auch der Intonationsverlauf zu den so genannten organisch bedingten Stimmmerkmalen.

4.3.3. Die Stimmqualität

Auch die Stimmqualität kann sprecherspezifische Merkmale aufweisen. Als prägnantes Merkmal ist die so genannte *Krächzstimme* (*creaky voice*) zu nennen. Zwar ist sie sowohl bei weiblichen als auch bei männlichen Sprechern weit verbreitet, jedoch gilt die Verteilung innerhalb der Rede als stark sprecherspezifisch. Meist tritt die Krächzstimme an Stellen auf, an denen die Stimme auf Grund des Intonationsverlaufes ohnehin absinkt, also am Ende von Sätzen oder anderen semantischen oder syntaktischen Einheiten (vgl. Künzel 1987:85, siehe auch Laver 1994:194-197, Baldwin & French 1990:57f). Sie kann jedoch bei Sprechern mit tiefer Stimmlage oder bei pathologischen Veränderungen der Stimmlippen (wie Laryngitis, Stimmlippenknötchen etc.) auch an unüblichen Stellen auftreten.

4.4. Individuelle Merkmale der Sprache

Im Bereich **Sprache** sind Dialekt, fremdsprachiger Akzent, Idiolekt (Summe individualtypischer Merkmale) und Soziolekt (Summe schichtspezifischer Merkmale, wie berufstypische Sprache / Jargon) von Relevanz (vgl. Jessen 2004, Sheffert et al. 2002:1448, Künzel 1998:385, 1997:104). Dabei gilt der *Dialekt* als „das mit Abstand bedeutendste sprecherspezifische Merkmal im Bereich der Sprache“ (Künzel 1987:85), das nicht nur die Aussprache sondern auch Teile der Grammatik und v.a. den Wortschatz betreffen kann. Da Dialekte oder Akzente in den meisten Fällen von sehr vielen Sprechern gesprochen werden, kommt den *idiolektalen Merkmalen* ein hoher individualtypischer Wert zu, da sie in der Regel nur für wenige Sprecher charakteristisch sind. Zu ihnen zählen neben sprechpathologischen Merkmalen (wie z.B. Sigmatismus, Schetismus, Balbuties etc.) auch nicht-pathologische idiolektale Merkmale wie Stereotypien und Hesitationsphänomene (vgl. Sheffert et al. 2002:1448, Baldwin & French 1990:53, Künzel 1986:86f). Auch *soziolektale Merkmale* geben Hinweise auf die Herkunft / Identität einer sprechenden Person. Hierbei ist insbesondere die verbale Ausdrucksfähigkeit (Eloquenz) von Bedeutung. Sie erlaubt u.U. Rückschlüsse auf Bildungsgrad und Berufsausbildung bzw. -ausübung. Letzteres kann sich in Form von spezifischen Redewendungen oder Fachausdrücken manifestieren. All diese Merkmale sind aus nahe liegenden Gründen zu den so genannten erworbenen sprecherspezifischen Merkmalen zu zählen.

4.5. Individuelle Merkmale der Sprechweise

Die individuelle **Sprechweise** wird u.a. beeinflusst von der Intonation (vgl. Kapitel 4.3.2), der Artikulationsgeschwindigkeit (Sprechtempo) und dem Atemverhalten (Frequenz, Dauer) (vgl. Jessen 2004, Künzel 1998:385, 1997:105). Dabei hat sich (v.a. in der Forensik) die *Geschwindigkeit des Sprechvorgangs* als sprecherspezifisch erwiesen. So zeigte Goldman-Eisler (²1973) unter verschiedenen Sprechern deutliche Unterschiede hinsichtlich ihrer Artikulationsrate⁸, jedoch eine stabile Konstanz innerhalb eines Sprechers, die von äußeren Faktoren kaum beeinflusst wird (siehe dazu auch Langenmayr 1997:265, 269ff).

Meist fällt das *Atemverhalten* eines Sprechers im Rahmen alltagskommunikativer Äußerungen erst dann auf, wenn dessen Ein- und Ausatmung nicht an syntaktisch-semanticen Grenzen der Rede sondern an anderen Stellen auftritt. Dies kann z.B. bei pathologisch begründeter Kurzatmigkeit der Fall sein. In wenigen Fällen kön-

⁸Anzahl der Silben pro Sekunde nach Abzug der Pausen (vgl. Goldman-Eisler ²1973).

4. Suprasegmentalia in der Spontansprache

nen auch *Atemgeräusche* individualtypisch sein und als sprecherspezifisches Merkmal dienen. Da beide Merkmalstypen jedoch eher selten anzutreffen sind, werden sie an dieser Stelle nicht weiter berücksichtigt.

Wie die obigen Ausführungen gezeigt haben, sind es v.a. die durch die Grundfrequenz bzw. deren Verlauf determinierten **Stimmerkmale**, die die entscheidenden Informationen einer Sprecheridentität enthalten. Dazu zählen insbesondere die *mittlere Sprechstimmlage*, *der Intonationsverlauf* sowie die *Stimmqualität*. Auch die Bereiche der individuellen **Merkmale der Sprache** und der **Sprechweise** sind dabei, v.a. für die forensische Sprechererkennung, von Bedeutung. Im Alltag ist jedoch davon auszugehen, dass naive Hörer die individuellen Merkmale der Sprache und der Sprechweise eher in Ausnahmefällen nutzen, um einen Sprecher zu identifizieren bzw. zu erkennen. Dies bestätigen auch die Untersuchungen zur Sprechererkennung, die im folgenden Kapitel vorgestellt werden sollen.

5. Naive und professionelle Stimmerkennung

5.1. Naive Stimmerkennung

Nachdem nun die erkenntnisrelevanten Sprechermerkmale und die Problematik der intra- und interindividuellen Sprechervariabilität erläutert wurden, beschäftigt sich das folgende Kapitel mit der Sprecher- und Stimmerkennung, die uns, zumeist unbemerkt, unzählige Male im Alltag gelingt. Zu diesem Zweck wurde das Kapitel in zwei Teile geteilt. Der Fokus von Kapitel 5.1.1 liegt auf den Einflüssen, die die Auswahl der Untersuchungsmethoden auf die Ergebnisse von Stimmerkennungsaufgaben haben können. Im Mittelpunkt von Kapitel 5.1.2 steht dabei insbesondere die Frage, welche konkreten (psycho) akustischen Parameter ein Hörender im Alltag möglicherweise nutzt, um eine Stimme zu erkennen.

5.1.1. Einflüsse auf die Probandenleistung bei Sprechererkennungsaufgaben

Bei jeder Studie, die sich mit Stimmerkennung beschäftigt, müssen zahlreiche Entscheidungen bei der Wahl der Methode getroffen werden. Bricker & Pruzansky (1974) veranschaulichen dies mit Hilfe eines stark vereinfachten und reduzierten Schemas der Kommunikationskette (vgl. Abbildung 5.1).

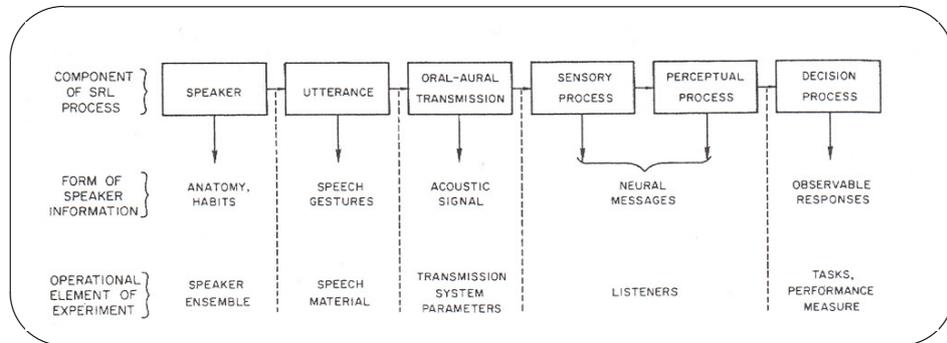


Abbildung 5.1.: *Speaker recognition by listening*: Schema mit Sprecherinformationen (aus: Bricker & Pruzansky 1974:298)

Im Rahmen dieses Unterkapitels und der einzelnen Abschnitte wird in aller Kürze gezeigt, welchen Einfluss die unterschiedlichen Variablen auf die Güte von Wiedererkennungseleistungen haben können.

Auswahl der Sprecherstimmen

Die relevantesten Entscheidungen bei der Auswahl der Sprecher betreffen Alter, Geschlecht und Dialekt der Sprechenden. Es ist nahe liegend, dass die Stimme einer Frau unter einer Auswahl männlicher Stimmen relativ leichter zu erkennen sein wird als unter einer Auswahl weiblicher Stimmen. Gleiches gilt auch für Alter (z.B. alt vs. jung) und Dialekt (z.B. rheinhessisch vs. ostwestfälisch).

Auch die Größe des Sprecherensembles nimmt Einfluss auf die Erkennungsrate. Wählt man die Methode der erzwungenen Auswahl unter mehreren Alternativen (x-alternate forced choice, X-AFC; nach Green & Swets ³1988), so sinkt die Wiedererkennungsrate mit einem Anstieg der Distraktorstimmenzahl. Dies bestätigten auch Legge et al. (1984). Bull & Clifford (²1987) zeigten, dass die Wiedererkennungsrate in einer 5-AFC-Aufgabe von 73 % auf 58 % bei einer 7-AFC-Aufgabe sank. Auch das Wissen der TeilnehmerInnen darüber, dass eine Stimme in einer bestimmten Auswahl vorkommen *muss*, beeinflusst die Wiedererkennung positiv (vgl. Bull, Rathborn & Clifford 1983, Warnick & Sanders 1980, Tosi et al. 1972).

Bei AFC-Aufgaben kann man außerdem zwischen *closed* und *open sets* unterscheiden. Bei *closed sets* sind die Teilnehmer über die Identität der Zielstimme informiert und wissen, dass sie sich unter den Stimuli befinden *muss* (vgl. Ladefoged & Ladefoged 1980). Bei Aufgaben mit *open sets* ist es möglich, dass sich die Zielstimme nicht zwangsläufig unter den präsentierten Stimmustern befinden muss. Warnick & Sanders (1980) konnten eine Zunahme korrekter Identifizierungen belegen, wenn die Teilnehmer nicht dazu gezwungen wurden, sich für eines der Stimmuster zu entscheiden. Dies bestätigen auch die Untersuchungen von Tosi et al. (1972).

Die wenigsten der bisher veröffentlichten Untersuchungen verwenden persönlich bekannte Stimmen als Sprecherstimmen. Wenn doch, so handelt es sich in der Regel um Arbeits- oder Studienkollegen. So zeigt ein früherer Überblick von Bricker & Pruzansky (1974), dass nur 8 von 57 Studien familiär bekannte Stimmen nutzen. Die anderen Studien bedienen sich fremder Stimmen, die vor der eigentlichen Wiedererkennungsaufgabe zunächst unter kontrollierten oder unkontrollierten Bedingungen trainiert werden müssen. Dies gilt auch für die Untersuchungen der jüngeren Vergangenheit (vgl. z.B. Imaizumi et al. 1997, Neuner & Schweinberger 2000, Nakamura et al. 2001). Legge, Grosman & Pieper (1984) zeigen, dass die Wiedererkennungsleistung mit der Anzahl der neu zu erlernenden Stimmen abnimmt. Gleiches gilt für die zeitliche Länge des Merkindervalls (vgl. u.a. Kerstholt et al. 2004, 2006, Nygaard &

5. Naive und professionelle Stimmerkennung

Pisoni 1998, Yarmey & Matthys 1992, Yarmey 1991, Thompson 1987, Legge, Grossmann & Pieper 1984, Clifford, Rathborn & Bull 1981, McGehee 1937, 1944). Andere Untersuchungen hingegen nutzen medial bekannte Stimmen, die den teilnehmenden Probanden (nach Selbstauskunft) gut bekannt sind (z.B. Van Lancker, Kreiman & Emmorey 1985, Van Lancker, Kreiman & Wickens 1985, Van Lancker et al. 1988).

Sprechmaterial und Stimuli

Zunächst erscheint es trivial, dass die Wiedererkennung eines Sprechers von der zeitlichen Länge eines Stimulus abhängig sein kann. Dennoch – oder vielleicht gerade deswegen – finden sich nur wenige Studien, die diese Frage zum Untersuchungsgegenstand haben. Pollack, Pickett & Sumby (1954) konfrontierten ihre Probanden mit Stimuli von 200 bis 3000 ms Länge. Sie konnten zeigen, dass die Wiedererkennungsrate in Abhängigkeit von der Stimulidauer bis zu einer Länge von 1000 ms stetig zunahm. Wurden die Stimuli von den zu erkennenden Sprechern geflüstert, so betrug der Zeitbedarf für eine korrekte Identifizierung durchschnittlich 3000 ms. Die Autoren äußern die Vermutung, dass es bei der Wiedererkennung jedoch weniger auf die zeitliche Dauer als auf das enthaltene Sprecherrepertoire ankommt (Pollack, Pickett & Sumby 1954:406f). Bricker & Pruzansky (1966) überprüften diese Hypothese, in dem sie ihre Probanden mit Stimuli unterschiedlicher Kategorien konfrontierten: (a) Satz, (b) zwei Silben, (c) einzelne Silbe, (d) CV-Cluster und (e) Vokal. Da die zeitliche Dauer der Ein- bzw. Zweisilber (beide je 117 ms) respektive die der CV- und V-Stimuli (446 bzw. 498 ms) in etwa gleich waren, sich die Wiedererkennungsrate bei diesen Stimuli jedoch unterschied (81 vs. 87 % bzw. 63 vs. 56 %), schlossen die Autoren daraus, dass die gestiegene Identifikationsleistung auf den vergrößerten Ausschnitt des Sprecherrepertoires zurückzuführen sein könnte (Bricker & Pruzansky 1966:1443). Compton (1963) nutzte im Rahmen seiner Untersuchung ausnahmslos den Vokal [i:] und bot ihn in unterschiedlicher zeitlicher Dauer an. Obwohl es sich lediglich um ein einzelnes Phonem und somit um einen winzigen Ausschnitt des Sprecherrepertoires handelt, konnte gezeigt werden, dass die Wiedererkennungsrate mit der Phonationsdauer zunahm (36 % bei 25 ms vs. 65 % bei 750 ms). Zu einem ähnlichen Ergebnis kamen Emmorey, Van Lancker & Kreiman (1984). Sie nutzten medial berühmte Stimmen und verwendeten drei unterschiedliche Stimulustypen: (a) spontane Rede (2.000 ms), (b) isolierte Wörter (durchschnittlich 481 ms) und (c) Vokalfolgen (durchschnittlich 494 ms). Bei den zeitlich längsten Stimuli lag die Erkennungsrate im Durchschnitt bei 61 %. Die weiteren Leistungen nahmen mit der Dauer der verwendeten Stimuli signifikant ab (40 % bei einzelnen Wörtern bzw. 34 % bei Vokalfolgen). Ein Einfluss der Phonemanzahl konnte dabei nicht nachgewiesen

5. Naive und professionelle Stimmerkennung

werden. Roebuck & Wilding (1993) konnten mit Hilfe einer 6-AFC-Aufgabe zeigen, dass in ihrer Stichprobe die Anzahl der enthaltenen Vokoide zu einer signifikanten Steigerung der Erkennungsrate führte (Roebuck & Wilding 1993:479f).

Da die bisherigen Untersuchungen widersprüchlich sind und sich kein einheitliches Bild zeigt, kann eine abschließende Antwort auf den Einfluss der Stimulidauer noch nicht gegeben werden. Es zeigt sich jedoch übereinstimmend, dass ein Zuhörer in der Regel von einer schrittweisen Verlängerung der Stimuli bis 1000 ms profitiert (vgl. u.a. Schweinberger, Herholz & Sommer 1997, Orchard & Yarmey 1995, Yarmey & Matthys 1992, Bricker & Pruzansky 1966).

Auch die Sprache, in der Stimuluswörter oder -sätze gesprochen werden, kann einen Einfluss auf die Wiedererkennung eines Sprechers haben. Doty (1998) zeigt beispielsweise, dass Stimmen, die in einer vertrauten Sprache sprechen, besser erkannt werden, als wenn sie eine fremde Sprache, oder akzentuiert sprechen (vgl. auch Goggin et al. 1991, Thompson 1987, Goldstein et al. 1981).

Schall- / Signalübertragung

Obwohl in den meisten Publikationen keine exakten Angaben über die verwendeten technischen Systeme und Speichermedien gemacht werden, soll an dieser Stelle auf die Unterschiede einzelner Verfahren hingewiesen werden.

McGehee nutzte während ihrer ersten Untersuchungen Ende der dreißiger Jahre die wohl verlustfreieste aller möglichen Methoden: sie separierte Sprecher und Hörer visuell voneinander und verzichtete auf sämtliche elektronische Speicher- oder Übertragungsmedien.

Generell spiegeln die in den bisherigen Studien verwandten Methoden der Schallübertragung und -speicherung den jeweiligen Stand des technisch Machbaren wider. Von einfachen *Spulentonbändern*, über *Kompaktkassetten* bis zur digitalen Aufnahmetechnik mit *MiniDisc* oder *DAT* sind alle Verfahren und technischen Methoden vertreten.

Bei den aktuellen Speicherverfahren kann sich aus technischen Gründen die so genannte Datenreduktion auf einzelne akustische Parameter auswirken. Dies bleibt für die Sprechererkennung zwar i.d.R. ohne Folgen, beeinträchtigt aber die phonetische Analyse hoher Frequenzen im Sprechschall (> 20-25 kHz). So nutzen z.B. MiniDisc-Recorder seit 1992 das *Adaptive Transform Acoustic Coding (ATAC)*. Dabei handelt

5. Naive und professionelle Stimmerkennung

es sich um eine Audioreduktion¹, die digitalisierte Sprachsignale in ihrer Größe reduziert. Die meisten Algorithmen versuchen dabei, das Signal so zu reduzieren, dass es für den Hörer ohne psychoakustische Folgen bleibt. So werden z.B. Frequenzen oberhalb von 20 (bis 25) kHz meist nicht mehr wahrgenommen und daher von den Systemen gar nicht erst aufgezeichnet, ebenso werden die schwer wahrnehmbaren, leisen Nebentöne vom Kompressionssystem entfernt. Das inzwischen weit verbreitete *ISO MPEG Audio Layer 3 (MP3)* reduziert das Datenvolumen bis maximal um den Faktor 50; eine Reduzierung um den Faktor 10 ist jedoch anhand von Qualitätseinbußen bereits wahrnehmbar. Aus diesen Gründen ist von einer Benutzung solcher verlustbehafteter Reduktion- oder Kompressionssysteme abzuraten.

Zu den verlustfreien Verfahren gehören – neben der Arbeit mit DAT-Bändern – u.a. *Apple Lossless Audio Codec (ALAC)*, *Free Lossless Audio Codec (FLAC)* oder *Lossless Predictive Audio Compression (LPAC)*. Apple Lossless ermöglicht es beispielsweise mit Hilfe der Software *iTunes* unkomprimiertes Audiomaterial (wie z.B. aus RIFF WAV-Dateien (.wav)) verlustfrei zu komprimieren.

Ende der achtziger Jahre löste das *Digital Audio Tape (DAT)*² die herkömmliche Kompaktkassette ab und zeigte eine deutliche Steigerung in Qualität und Komfort. Mit einer Samplingfrequenz von bis zu 96 kHz und einer Auflösung von maximal 24 Bit verfügt es heute über ein höheres Klangpotenzial als die *Compact Disc (CD)* und bietet zudem eine absolut verlustfreie Datenreduktion. Aus diesem Grund ist das DAT-Band bei aktuellen (phonetischen) Untersuchungen häufig Mittel der Wahl.

Bei der Aufzeichnung von Stimmsignalen ist also darauf zu achten, dass nur Methoden verwandt werden, die eine möglichst geringe, am besten überhaupt keine verlustbehaftete Datenreduktion zur Folge haben. Die wenigen Untersuchungen zur Stimmerkennung in Abhängigkeit der Übertragungsbedingungen stammen aus dem Bereich der Forensik und beschäftigen sich hauptsächlich mit der telefonischen Signalübermittlung (siehe z.B. Künzel 1990:22-38; siehe auch Kapitel 5.2).

Auch zu den verwendeten Lautsprechersystemen gibt es in der Regel keine weiteren Auskünfte. Hierbei gilt zu beachten, dass die Lautsprecher alle Frequenzen zwischen 20 bis mindestens 25000 Hz gleich laut wiedergeben. Bei einigen Geräten kommt es zu so genannten *linearen Verzerrungen*, was bedeutet, dass unterschiedliche Frequenzen trotz eines identischen Eingangssignalpegels verschieden laut wiedergegeben werden. Bei handelsüblichen Geräten sind Abweichungen bis $\pm 0,5$ dB unhörbar und somit

¹Mitunter findet auch der Begriff *Audiokompression* Verwendung, der nicht ganz korrekt ist, da es in der Regel zu einer verlustbehafteten Datenreduktion kommt.

²Streng genommen gab es zwei unterschiedliche Verfahren – S-DAT und R-DAT – wobei sich letzteres aus unterschiedlichen Gründen bei den Privatanutzern jedoch nicht durchsetzen konnte.

5. Naive und professionelle Stimmerkennung

als unbedenklich anzusehen.

Hörerensemble

Eine große Rolle spielt die Zusammensetzung des Hörerensembles. Der größte Teil der bisherigen Studien beschäftigt sich mit den Wiedererkennungsleistungen uneinträchtigter, also „gesunder“ Erwachsener. Wie bereits in dem Abschnitt *Sprechmaterial und Stimuli* erläutert, hängt die Wiedererkennungsleistung der Probanden in großem Umfang von den jeweils gewählten Untersuchungsmethoden ab. Ohne an dieser Stelle näher auf die einzelnen Methoden eingehen zu können, zeigen die Leistungen der (zerebral) gesunden Normalpersonen eine erhebliche Normvarianz. Ausgewählte Untersuchungsergebnisse sind in Tabelle 5.1 zusammengefasst. Da die reinen Wahrnehmungs- bzw. Erkennungsexperimente in den letzten Jahren zugunsten von Forschungen mit Hilfe von EEG- und fMRT deutlich abgenommen haben, finden sich dort überwiegend ältere Untersuchungen.

5. Naive und professionelle Stimmerkennung

| Autoren | Anzahl Stimmen | Anzahl Hörer | Stimuli | percent correct | Streuung |
|---------------------------------------|----------------|--------------|---------------------|-----------------|-----------|
| Bricker & Pruzansky (1966) | 10 | 16 | Vokale | 55 | 39-75 |
| | | | CVC-Cluster | 61,5 | 47-83 |
| | | | Einsilber | 83,5 | 64-94 |
| | | | Zweisilber | 88 | 73-98 |
| Van Lancker, Kreiman & Emmorey (1985) | 45 | 72 | spontan (2s) | 17,4 | 0-71,1 |
| | | | w.o. (6-AFC) | 14,57 | 26,7-91,1 |
| | | | spontan (4s rückw.) | 13,52 | 15,6-75,6 |
| | | | spontan (2s) | 26,6 | 0-80 |
| Van Lancker, Kreiman & Wickens (1985) | 30 | 72 | „known“, max. 45 | 69,9 | 40-93,5 |
| | | | w.o. (6-AFC) | 57,5 | 19,2-92,3 |
| | | | spontan (4s rückw.) | 67,91 | 20-100 |
| | | | um 33 % gedehnt | 56,11 | 20-90 |
| Van Lancker & Kreiman (1987) | 25 | 48 | um 33 % gestaucht | 57,78 | 10-100 |
| | | | spontan (2s) | 75,38 | 33,3-100 |
| | | | um 33 % gedehnt | 59,4 | 12,5-100 |
| | | | um 33 % gestaucht | 63,12 | 14,3-100 |
| Neuner & Schweinberger (2000) | 30 | 20 | spontan (4s, 4-AFC) | 82,1 | 46,7-100 |
| | | | spontan (2s) | 43,8 | 37-50 |
| Nakamura et al. (2001) | >50 | 9 | spontan (2s) | 84 | 68-97 |

Anm.: Angaben bei *percent correct* und Streuung in %; „known“ = Stimmen, von denen der Proband glaubt, sie erkennen zu können.

Tabelle 5.1.: Streuung der Wiedererkennungsrates gesunder Probanden in ausgewählten Untersuchungen

5. Naive und professionelle Stimmerkennung

Wie die Streuungswerte in Tabelle 5.1 zeigen, gibt es offenbar Menschen, denen es leicht fällt, eine Stimme zu erkennen, während hingegen andere dabei große Probleme zeigen. In ihrer Studie von 1937 befragte McGehee im Anschluss die teilnehmenden Probanden nach deren individuellen Einschätzung der gehörten Stimmen. Die Antworten reichten von: „The test was not fair because the voices were too dissimilar, so that it was easy to pick out the repeating voice“ bis „It would not surprise me if they were all made by the same person“ (McGehee 1937:261). Auch Abberton & Fourcin postulierten nach der Auswertung ihrer Ergebnisse: „Subjects vary significantly in their ability to identify speakers“ (Abberton & Fourcin 1978:314).

Einige Untersuchungen beschäftigten sich u.a. mit der Frage nach etwaigen Geschlechtsunterschieden bei der Wiedererkennung von Stimmen. So zeigten Wilding & Cook (2000), dass weibliche Probanden im Rahmen einer 6-AFC-Aufgabe die Stimmen ihrer Geschlechtsgenossinnen signifikant besser erinnerten als die der männlichen Sprecher (51 % vs. 43 %). Umgekehrt war dies bei den männlichen Teilnehmern nicht der Fall (vgl. Wilding & Cook 2000:535). Zu anderen Ergebnissen kam die Untersuchung von Günzburger (1984:15). Hier zeigten die weiblichen Probanden eine bessere Erkennungsrate bei männlichen als bei weiblichen Stimmen. Auch die männlichen Teilnehmer erinnerten die weiblichen Stimmen besser als die ihrer Geschlechtsgenossen.

Unabhängig vom Geschlecht der Sprecher, zeigte die Untersuchung von Van Lancker, Kreiman & Emmorey (1985) eine tendenziell, aber nicht signifikant bessere Leistung der männlichen gegenüber der weiblichen Teilnehmer („normal condition“: Männer 71,5 % vs. 68,5 %; „backward“: Männer 59,3 % vs. 55,9 %, vgl. Van Lancker, Kreiman & Emmorey 1985:28). Dies bestätigen auch Sheffert et al. (2002:1452). Auch der Faktor „Alter“ zeigte keine signifikanten Korrelationen mit den erbrachten Leistungen (vgl. Van Lancker, Kreiman & Emmorey 1985:29). Zu anderen Ergebnissen kommt Künzel (1990). Im Rahmen seiner Stichprobe konnte er einen signifikanten Zusammenhang zwischen Alter der Probanden und ihrer Fehlerhäufigkeit nachweisen, wobei die Zunahme lediglich die Falsch-Identifizierungen betraf und nicht die Falsch-Zurückweisungen. Einen Geschlechtereffekt konnte Künzel nicht nachweisen (vgl. Künzel 1990:54f, 58).

Aufgrund dieser uneinheitlichen und zum Teil auch widersprüchlichen Untersuchungsergebnisse ist zum jetzigen Zeitpunkt keine abschließende Aussage über den Einfluss von Geschlechtszugehörigkeit und / oder Alter möglich.

5. Naive und professionelle Stimmerkennung

Wie bereits oben erwähnt, beschäftigen sich die meisten Untersuchungen zur Sprechererkennung mit den Leistungen Erwachsener. Einige Untersuchungen, die mit Kindern durchgeführt wurden, sollen im Folgenden kurz vorgestellt werden.

Die Frage, ob bereits Föten im Mutterleib in der Lage sind Stimmen zu unterscheiden, untersuchten u.a. Kisilevsky et al. (2003). Dazu wurden 60 Föten im Alter von durchschnittlich 38,4 Wochen mit rund 2-minütigen Textpassagen beschallt. 30 Föten präsentierte man einen Text, den die jeweilige Mutter gesprochen hatte, die 30 anderen Föten „hörten“ denselben Text von einer fremden Frau. Die Gruppe, der die vertraute Stimme präsentiert wurde, zeigte 26 s post-onset bis zum Stimulusende eine signifikant höhere Herzfrequenz als die Föten aus der anderen Gruppe (vgl. Kisilevsky 2003:222). Ein signifikanter Anstieg der Körperbewegungen konnte nicht gezeigt werden. Dies bestätigen auch die Ergebnisse der Studie von Hepper, Scott & Shahidullah (1993). Eine Untersuchung von DeCasper & Fifer (1980) zeigt, dass auch Neugeborene im Alter von drei Tagen auf die Stimme ihrer Mutter anders reagieren als auf die Stimme einer fremden Frau (vgl. DeCasper & Fifer 1980:1174). Ähnliches berichten Roe (1978) und Turnure (1971) für Säuglinge im Alter von drei bis neun Monaten. Mann, Diamond & Carey (1979) überprüften die Fähigkeit von Grundschulern, eine gelernte Stimme wiederzuerkennen. Dabei wiesen sie eine Korrelation zwischen Alter und Wiedererkennungsrate nach. Eine Untersuchung von Bennet & Montero-Diaz (1982) ergab, dass 6- bis 9jährige Kinder das Geschlecht eines Sprechers ebenso zuverlässig anhand einer Stimmprobe erkennen konnten wie Erwachsene (vgl. Bennet & Montero-Diaz 1982). Eine ähnliche wie die bereits beschriebene Studie von Van Lancker, Kreiman & Emmorey (1985) wurde von Bartholomeus (1973) mit Kindern zwischen vier und fünf Jahren durchgeführt. In dieser Studie bestanden die 18 Stimuli aus je 3-sekündigen Sprechproben von Kindern aus der gleichen Kindergartengruppe. Wie bei den Erwachsenen zeigte sich auch hier eine geringere Wiedererkennungsrate der rückwärts präsentierten Stimuli. Außerdem konnte eine erhebliche Normvarianz bei der individuellen Erkennungsleistung gezeigt werden. Die Streuung reichte bei der „voice-picture matching task“ von 11 bis 100 %, bei der „voice-naming task“ von 12 bis 95 % (vgl. Bartholomeus 1973:470), wobei die Mittelwerte unterhalb derer gesunder Erwachsener rangierten. Aus der Studie geht jedoch nicht hervor, wie vertraut die Kinder mit den Stimmen ihrer Kameraden gewesen sind.

Spence, Rollins & Jerger (1997) untersuchten die Wiedererkennungsfähigkeit von 27 Vorschulkindern im Alter zwischen drei und fünf Jahren mit Hilfe der Stimmen bekannter Comic-Figuren (z.B. Kermit, Miss Piggy, Ernie, Bugs Bunny, Krümel-

5. Naive und professionelle Stimmerkennung

monster, Fred Feuerstein etc.). Im Rahmen eines voice-picture-Paradigmas wurden die Kinder gebeten, zwanzig 4-sekündige Stimmproben jeweils einer von sechs Comic-Figuren zuzuordnen (6-AFC). Konnten die Kindern den Namen der Figur nennen, so wurde die Stimme als „*more familiar*“ eingestuft, formulierten sie Äußerungen wie „*Der ist von der Sesamstraße*“, so wurde sie als „*less familiar*“ eingestuft (vgl. Spence, Rollins & Jerger 1997:219). Die Gruppe der Dreijährigen (36 - 37 Monate) erkannte (unabhängig vom Bekanntheitsgrad *more vs. less familiar*) durchschnittlich 61,44 %, die Vierjährigen (48 - 59 Monate) 81,36 % und die Fünfjährigen (60-71 Monate) 86,03 %. Die Ergebnisse ergaben zwei signifikante Haupteffekte: das Alter der Kinder und, übereinstimmend mit den Untersuchungen Erwachsener, die Vertrautheit mit der zu erkennenden Stimme. Da jedoch die Stimmen der verwendeten Comic-Figuren eine wesentlich größere Heterogenität zeigen als menschliche Stimmen, lassen sich die Ergebnisse nur eingeschränkt mit denen von Erwachsenen vergleichen.

Neben den Untersuchungen mit gesunden Erwachsenen und Kindern existieren darüber hinaus einige wenige Studien, die sich mit den Leistungen von beeinträchtigten Personen beschäftigen. Eine Auswahl zeigt Tabelle 5.2. Auf Untersuchungen, die sich speziell mit rechts- bzw- linkshirngeschädigten Patienten beschäftigen, wird ausführlich in Kapitel 7 eingegangen.

| Autor/en | Beeinträchtigung | N | Ergebnis |
|-------------------------|----------------------|----|-----------------------------------|
| Bull et al. (1983) | Sehbeeintr. | 92 | signifik. besser als KG |
| Reid et al. (1993) | Blinder Capgras-Pat. | 1 | „deficit“ |
| Brachacki et al. (1994) | Dyslexie | 7 | signifik. schlechter als KG |
| Boucher et al. (1998) | Autismus | 7 | signifik. schlechter als KG (SLI) |
| Boucher et al. (2000) | Autismus | 19 | kein Unterschied zur KG (SLI) |

Tabelle 5.2.: Übersicht über Stimmerkennungsstudien mit beeinträchtigten Probanden

In der Studie von Bull, Rathborn & Clifford wurden 92 Probanden im Alter von 16 bis 52 Jahren gebeten, im Rahmen einer 5-, 7- oder 9-AFC-Aufgabe eine fünf Sekunden zuvor einmalig gehörte Stimme wiederzuerkennen. Jede Aufgabe wurde sechs mal mit unterschiedlichen Ziel- und Distraktorstimmen durchgeführt. Als Stimulus diente ein etwa viersekündiger Satz („*I'll meet you outside the Westminster Bank at six o'clock tonight.*“, Bull, Rathborn & Clifford 1983:224). Das Sehvermö-

5. Naive und professionelle Stimmerkennung

gen der Probanden fehlte völlig oder war herabgesetzt („*totally blind; perception of light; residual/goodish sight*“, Bull, Rathborn & Clifford 1983:224). Als Kontrollgruppe dienten 72 normalsehende Personen. Die Analyse der Ergebnisse erbrachte einen signifikanten Unterschied zwischen den Leistungen der sehbeeinträchtigten und der normalsehenden Teilnehmer ($p < ,0025$). Zwischen den Ergebnissen der 5-, 7- bzw. 9-AFC-Aufgaben zeigten sich keine signifikanten Unterschiede. Weder der Grad der Sehbeeinträchtigung, noch das Alter der Erblindung oder der Intelligenzquotient korrelierte mit der Erkennungsrate.

Brachacki, Fawcett & Nicolson (1994) untersuchten sieben dyslexische Studenten und acht Kontrollprobanden hinsichtlich ihrer Fähigkeit der Sprecheridentifizierung. Zu diesem Zweck präsentierte man ihnen für jeweils drei Sekunden acht Gesichter. Parallel wurde die entsprechende Stimme eingespielt. Alle Stimmen sprachen dabei den gleichen Satz („*Hello, I am a student at Sheffield University, and my name is X*“, Brachacki, Fawcett & Nicolson 1994:305). Die acht bekannten Stimmen wurden um acht weitere unbekannte Distraktorstimmen ergänzt und in randomisierter Reihenfolge – unter Löschung des Namens der sprechenden Person – präsentiert. Die gleiche Prozedur wurde mit den Gesichtern wiederholt. Nach einer Woche wurden alle Probanden ein zweites Mal getestet. Die Ergebnisse sind in Tabelle 5.3 aufgeführt (vgl. Brachacki et al. 1994:306).

| Test | Gruppe | Stimme | | Gesichter | |
|----------|----------------|--------|------|-----------|------|
| | | M | SD | M | SD |
| Ersttest | Dyslektiker | 58,9 | 12,0 | 98,2 | 4,7 |
| | Kontrollgruppe | 78,1 | 13,2 | 98,5 | 4,4 |
| Re-Test | Dyslektiker | 51,8 | 9,0 | 83,3 | 19,1 |
| | Kontrollgruppe | 60,9 | 15,6 | 95,3 | 5,3 |

Anm.: alle Angaben in %.

Tabelle 5.3.: Leistungen der dyslektischen und der Kontrollprobanden

Die Analyse der Daten zeigte einen signifikanten Unterschied der Stimmerkennungsleistungen zwischen der Dyslexie- und der Kontrollgruppe ($p < ,05$), bei der Gesichtererkennung ($p > ,1$) jedoch nicht. Trotz einer, durch die geringe Probandenzahl verursachten, eingeschränkten Aussagekraft der Ergebnisse, geben sie dennoch

5. Naive und professionelle Stimmerkennung

einen interessanten Hinweis auf die Defizite im Bereich der auditiven Verarbeitung dyslektischer Erwachsener, die sich mit den Annahmen über die defizitäre phonologische Bewusstheit im weiteren Sinne (vgl. Marx 1997) decken könnten. Eine Studie zur Stimmerkennung lese- und / oder rechtschreibbeeinträchtigter Kinder ließe aufschlussreiche Ergebnisse erwarten.

In gleich zwei Studien beschäftigten sich Boucher, Lewis & Collis (1998, 2000) u.a. mit der Fähigkeit autistischer Kinder, die Stimmen ihrer Lehrer bzw. anderer Erwachsener ihrer Schule zu identifizieren. Nach einer Trainingsphase wurden den Kinder im Alter von 7;10 (Autisten, N = 7) bzw. 7;9 Jahren (Kontrollgruppe mit Lernbehinderung und SLI³, N = 8) insgesamt 30 Testsätzen („*Hello. Do you know who I am? Am I at your school?*“, vgl. Boucher, Lewis & Collis 1998:177) präsentiert, von denen ihnen 15 Sprecher vertraut und weitere 15 Sprecher unbekannt waren. Anhand von Fotografien sollten die Probanden die Stimmen den entsprechenden (Lehr-) Personen zuordnen. Eine Analyse der Messwerte zeigte, dass die Leistungen der autistischen Kinder signifikant schwächer waren als die der Kontrollgruppe ($p = .04$). Für ihre zweite Studie (Boucher, Lewis & Collis 2000) akquirierten sie u.a. 19 Kinder mit Autismus (9;7 Jahre) und weitere 19 Kinder mit SLI (9;0 Jahre). Als Stimuli dienten auch hier die Stimmen von Lehrern und anderen Mitarbeitern der Schule sowie von Klassenkameraden. Nach zwei Trainingsphasen wurden die Teilnehmer u.a. mit 18 Stimmmustern konfrontiert (neun bekannt, neun unbekannt). Bei den Teststimuli handelte es sich um Ausschnitte aus Spontansprachstichproben, deren exakte Länge nicht angegeben wurde. Die Analyse der Wiedererkennungsrates zeigte keinen signifikanten Unterschied zwischen den Leistungen der autistischen Kinder und denen der SLI-Gruppe. Aufgrund dieser, sich widersprechenden Ergebnisse, ist eine abschließende Aussage über die Fähigkeit der Stimmentifizierung bei Kindern mit Autismus nicht möglich. Ein Grund für die unterschiedlichen Ergebnisse mag darin liegen, dass – obwohl genaue Zeitangaben fehlen (s.o.) – die Stimulilänge und der Ausschnitt des entsprechenden Sprecherrepertoires erheblich größer war als bei der Studie von 1998 („*All the recordings were of people talking about going shopping or what they will do if they won the lottery*“, Boucher, Lewis & Collis 2000:851 vs. „*Hello. Do you know who I am? Am I at your school?*“, vgl. Boucher, Lewis & Collis 1998:177).

³*Specific Language Impairment.*

5. Naive und professionelle Stimmerkennung

Anhand unterschiedlicher Wahrnehmungsexperimente mit sehbeeinträchtigten, autistischen oder dyslektischen Personen konnte gezeigt werden, dass die Fähigkeit zur Sprecheridentifizierung bzw. -erkennung von ganz unterschiedlichen pathologischen Faktoren beeinflusst werden kann.

Das Wiedererkennen von Stimmen ist darüber hinaus keine rein menschliche Fähigkeit. So konnte Insley (2000) bei einer Population Nördlicher Seebären (*Callorhinus ursinus*) nachweisen, dass diese auch nach vier Jahren noch in der Lage war, die Vokalisationen ihrer Artgenossen zu erkennen (vgl. Insley 2000:404). Eine ähnliche Leistung beobachteten Charrier, Mathevon & Jouventin (2001, 2002) bei Muttertieren des Subantarktischen Seebärs (*Arctocephalus tropicalis*). Während der Futtersuche müssen die Muttertiere ihre Jungen zurücklassen und bei ihrer Rückkehr unter hundert anderen wiederfinden. Dies gelingt ihnen durch die Rufe der Jungtiere (vgl. Charrier, Mathevon & Jouventin 2001:603, 2002:873).

5.1.2. Sprechererkennung im Alltag

Von besonderer Bedeutung für die Forschungen im Bereich der Sprechererkennung ist die Bestimmung jener Parameter, die von einem Hörer genutzt werden, um einen Sprecher / eine Stimme zu identifizieren bzw. zu erkennen.

Frühe Studien zur naiven Sprechererkennung bedienten sich zunächst einer Methodenkombination bestehend aus Rating-Skalen auf Grundlage wahrgenommener, subjektiver also psychoakustischer Phänomene und objektiver physikalischer Messungen (vgl. z.B. McGehee 1944, Osgood, Suci & Tannenbaum 1957, Voiers 1964, Holmgren 1967). Holmgren (1967) gelang es mit Hilfe korrelativer Analysen semantischer Differenzialskalen zu zeigen, dass Hörer offenbar Merkmale wie die durchschnittliche Grundfrequenz, deren Variabilität, Amplitude und die Sprechgeschwindigkeit nutzen, um eine Stimme zu identifizieren. Dabei konnten *clarity* (Klarheit) / *intelligibility* (Verständlichkeit) am besten mit der *Amplitude* (Lautheit) und *pitch* (Tonhöhe) mit der Grundfrequenz F_0 beschrieben werden.

Eine weitere Möglichkeit zur Feststellung relevanter Parameter besteht darin, ausgewählte Variablen eines erzeugten Signals zu manipulieren. Dazu wurden nun erstmalig bekannte Stimmen, meist von Arbeitskollegen oder berühmten Persönlichkeiten, eingesetzt. Sanken die Erkennungsraten der Probanden infolge einer spezifischen Manipulation, so schrieb man dem veränderten Parameter eine entsprechende bedeu-

5. Naive und professionelle Stimmerkennung

tungstragende und somit eine für den Erkennungsprozess relevante Funktion zu. Es ist jedoch denkbar, dass die perzeptive Relevanz eines Parameters möglicherweise vom Vorhandensein oder Fehlen anderer Parameter abhängig sein kann. Dennoch bediente man sich dieser Methoden und nutzte bereits in den fünfziger Jahren elektronische Tiefpass- und Hochpassfilter, um bestimmte Frequenzbereiche zu eliminieren. Pollack, Pickett & Sumbly (1954:404f) veränderten beispielsweise monosilbische Stimuli mit Hilfe von Hoch- und Tiefpassfiltern, deren Cut-off-Frequenzen zwischen 100 und 5.000 Hz variierten. Sie zeigten, dass das Entfernen der spektralen Energie unterhalb von 500 Hz bzw. oberhalb von 3 kHz keinen Effekt auf die Sprechererkennungsraten hatte. Wurden die Frequenzen oberhalb von 500 Hz entfernt, so hatte dies jedoch ein deutliches Absinken der Erkennungsrate zur Folge (Skalbeck 1955, zit. n. Hecker 1971:40). Daraus konnte zunächst geschlossen werden, dass sich die relevanten, frequenzabhängigen Informationen in einem Bereich zwischen 0,5 und 3 kHz bewegen sollten.

In der Mitte der siebziger Jahre ging man davon aus, dass in der alltäglichen Sprechererkennung die glottale Signalquelle („*glottal source*“, Bricker & Pruzansky 1974:313), also die Tätigkeit der Stimmlippen („*vocal fold function*“, Colton & Estill 1979:115) für die Individualität einer Stimme von größerer Bedeutung ist, als die Signalmodifikation durch die Anatomie des Vokaltraktes (vgl. Bricker & Pruzansky 1974:313, Hollien 1990:196, siehe auch Seite 66). Erstere beeinflusst vor allen Dingen die mittlere Grundfrequenz (*fundamental mean frequency*) und deren Verlauf (*variance*). Künzel (1987) formuliert, dass diese „*das wichtigste Merkmal bei der Sprechererkennung darstellt*“ (Künzel 1987:82, Hervorhebung im Original; vgl. auch Rose 2002:244, Hollien 1990:196, Baldwin & French 1990:48, Nolan 1983:124, Abberton & Fourcin 1978:314, Walden et al. 1978, Scherer 1974).

Eine Untersuchung, die diese Aussage zumindest zum Teil unterstützt, stammt von Van Lancker, Kreiman & Emmorey (1985). Sie untersuchten im Rahmen ihrer Studie die Wiedererkennungsraten von 94 Probanden mit Hilfe der Stimmen berühmter Sprecher (64 männliche Entertainer, Politiker etc.)⁴. Die Untersuchung gliederte sich in drei Teile. Dabei bestanden die Stimuli in

- Task 1 aus den Stimmustern von jeweils 2 s Länge in einem *open set*⁵,
- Task 2 ebenfalls aus Stimmustern von jeweils 2 s Länge in einem *closed set*⁶

⁴U.a. Steve Martin, John F. Kennedy, Bob Hope, Ronald Reagan (vgl. Van Lancker, Kreiman & Emmorey 1985:51).

⁵In einem *open set* bleibt offen, ob die Zielstimme unter den Stimuli vorhanden ist oder nicht (vgl. Seite 51).

⁶In einem *closed set* werden die Probanden i.d.R. über die Identität des Sprechers informiert und

5. Naive und professionelle Stimmerkennung

mit sechs Antwortmöglichkeiten (6-AFC) und

- Task 3 aus Stimmproben von 4 s Länge, die rückwärts abgespielt wurden (ebenefalls in einem *closed set* mit sechs möglichen Antworten, 6-AFC).

Die Probanden wurden in insgesamt vier Gruppen eingeteilt. Die Gruppen I-III wurden nach Alter zusammengesetzt und absolvierten Task 1, 2 und 3. Gruppe IV setzte sich aus 20 Personen zwischen 13 und 44 Jahren zusammen und absolvierte nur Task 3 und 2. Die mittlere Erkennungsleistung bei Task 1 betrug 26,6 %, bei Task 2 69,9 % und bei Task 3 57,5 %. Das Ergebnis von Task 1 (26,6 %) entspricht in etwa dem Ergebnis von Ladefoged & Ladefoged (1980), die ein ähnliches Design verwendeten. Auch das Ergebnis von Task 2 bestätigt die bereits von anderen Autoren vorgenommenen Untersuchungen der Stimmerkennung im Rahmen von *closed sets* (AFC 5-10, 5 - 10 Antwortmöglichkeiten) (vgl. z.B. Hollien, Majewski & Doherty 1982, Ladefoged & Ladefoged 1980, Bricker & Pruzansky 1966 etc.). Die hohe Sprechererkennungsrate bei Task 3 zeigt, dass ein Erkennen auch dann möglich ist, wenn sich die akustische Information auf *pitch*, *pitch range*, Geschwindigkeit, Stimm- und Vokalqualität beschränkt (vgl. Van Lancker, Kreiman & Emmorey 1985:30). Eine Aufschlüsselung der erfolgreich erkannten Stimmen zeigt, dass manche Stimmen vor- und rückwärts⁷ etwa gleich gut erkannt werden konnten, während andere vorwärts deutlich besser erkannt wurden als rückwärts (vgl. Tabelle 5.4). Bei einigen Stimmen trat auch der umgekehrte Fall ein: sie konnten rückwärts zuverlässiger erkannt werden als vorwärts (vgl. Van Lancker, Kreiman & Emmorey 1985:36). Dies zeigt, dass die durch das rückwärts Spielen beeinflussten Artikulationsmuster (*specific articulatory patterns*) für das Erkennen mancher Stimmen essenziell sind, während sie für andere Stimmen irrelevant zu sein scheinen.

In einer weiteren Studie untersuchten Van Lancker, Kreiman & Wickens (1985) nach einem ähnlichen Design wie eben beschrieben die Rolle der Sprechgeschwindigkeit. Dazu nutzten sie die 30 Stimmen, die in der vorhergehenden Studie (Van Lancker, Kreiman & Emmorey 1985) am besten erkannt wurden. Jedes der Stimmuster wurde im Rahmen der Untersuchung in drei Varianten als Stimulus verwendet: (a) im Original, (b) gedehnt (um 33 %, 5320 ms), (c) komprimiert (um 33 %, 2680 ms). Die Grundfrequenz änderte sich durch die Veränderung der Geschwindigkeit nicht. Die Probanden wurden in drei Gruppen aufgeteilt und jeder Teilnehmer hatte die Aufgabe, 30 unterschiedliche Stimuli (je zehn (a) im Original, (b) gedehnt, (c) kom-

wissen, dass sie im Stimulus enthalten ist (vgl. Seite 51).

⁷Dabei muss jedoch zusätzlich beachtet werden, dass die Vorwärts-Proben eine Länge von 2 s hatten, während die Rückwärts-Proben 4 s dauerten. Eine Vergleichbarkeit der Ergebnisse erscheint also fraglich.

5. Naive und professionelle Stimmerkennung

| Stimme | vorwärts (2s) | rückwärts (4s) |
|---------------|---------------|----------------|
| Jack Klugman | 44,3 | 41,0 |
| Steve Martin | 87,1 | 80,6 |
| Leonard Nimoy | 88,1 | 85,1 |
| Bob Hope | 78,0 | 28,6 |
| Lawrence Welk | 86,8 | 36,8 |
| Richard Nixon | 18,0 | 77,8 |
| Anthony Quinn | 27,5 | 45,0 |

Anm.: Angaben in %.

Tabelle 5.4.: Wiedererkennungsrates berühmter Stimmen (vor- vs. rückwärts)

primiert) im Rahmen eines *closed set* mit sechs Antwortmöglichkeiten zu erkennen (6-AFC). Auch hier zeigen die Ergebnisse (vgl. Tabelle 5.5), dass das Erkennen einiger Stimmen durch die Manipulation der Geschwindigkeit nur wenig oder gar nicht beeinflusst wurde. Andere hingegen wurden in der gedehnten Version bzw. in der komprimierten Version besser erkannt als im Original (vgl. Van Lancker, Kreiman & Wickens 1985:51).

| Stimme | Original | gedehnt | gestaucht |
|--------------------|----------|-------------|-------------|
| Everett Dirksen | 91,7 | 84,6 | 92,9 |
| W. C. Fields | 100 | 95,8 | 100 |
| Nelson Rockefeller | 54,5 | 86,7 | 50,0 |
| Jimmy Durante | 21,7 | 21,7 | 37,5 |

Anm.: Angaben in %.

Tabelle 5.5.: Wiedererkennungsrates berühmter Stimmen (unterschiedliche Tempi)

Da es sich bei einem Stimmsignal um ein hochkomplexes, akustisches Muster handelt, weisen mehrere Autoren übereinstimmend darauf hin, dass für das Erkennen einer Stimme in Alltagssituationen offenbar mehrere Faktoren von Relevanz sein müssen (vgl. u.a. Sheffert et al. 2002:1455, Kreiman et al. 1993:32, Brown 1981:295, Abberton & Fourcin 1978:314). Zu diesem Faktorenbündel gehören nach Hollien (1990:196):

5. Naive und professionelle Stimmerkennung

- die mittlere Sprechstimmlage,
- die Formantfrequenzen,
- sowie Verhältnis (ratio) und
- Transitionen der Vokoide.

Ergänzend formuliert er:

„There is little question but that these features provide important cues in the perceptual identification of speakers. As a matter of fact, attempts have been made (...) to compare the relative importance of source (voice) and vocal tract (articulatory system) transfer characteristics for speaker identification purposes; it was found that these features contribute additively“ (Hollien 1990:196).

Zusammenfassend lässt sich also sagen, dass es nicht *den* Faktor gibt, der allein für ein erfolgreiches Erkennen verantwortlich ist, denn dem liegt immer eine ausschließlich subjektive Leistung zugrunde. Entsprechend formulieren Van Lancker und Kollegen:

„This observation led us to suggest that a given parameter (...) is central to the characterization or 'quality' of some voices, but irrelevant to others, and that the primary or indispensable cue(s) to voice identity vary from voice to voice“ (Van Lancker, Kreiman & Wickens 1985:40).

Und an anderer Stelle:

„These and other results suggests that each familiar voice is best viewed as a relatively unique pattern, and that human voice recognition operates by utilizing different acoustic parameters for different voices“ (Van Lancker, Kreiman & Emmorey 1985:19).

Bei einer Beurteilung der Ähnlichkeit von Stimmen konnten Kreiman et al. (1990) deutliche Unterschiede innerhalb und zwischen den getesteten Gruppen nachweisen (z.B. geschulte Phonetiker vs. linguistisch Naive, vgl. Kreiman et al. 1992).

Entsprechend formulieren auch Van Lancker, Kreiman & Emmorey:

„(...) it is not useful to pursue *the* parameter that contributes most universally to voice identity. Instead, many parameters constitute a pool from which certain selected cues are utilized for recognition“ (Van Lancker, Kreiman & Emmorey 1985:33; Hervorhebung im Original).

5. Naive und professionelle Stimmerkennung

Dies wird deutlich und nachvollziehbar, wenn man das Phänomen der Stimmerkennung mit dem der Gesichtererkennung vergleicht. Auch hier können wir als handelnde Subjekte nur schwer feststellen, woran wir letztendlich das Gesicht einer Person erkannt haben: an der Augenfarbe, der Frisur, der Nase, dem Mund?

5.2. Professionelle Sprechererkennung in der Forensik

Wie in den vorangegangenen Kapiteln erläutert, beinhaltet des Stimmsignal einer sprechenden Person spezifische akustische Merkmale und Parameter, die für eine bestimmte Person individualtypisch und charakteristisch sind. Diese Merkmale macht sich die forensische Sprechererkennung bei der Aufklärung von Straftaten zu Nutze, bei denen ein Täter seine Stimme als Tatwerkzeug eingesetzt hat. Wie die Forensik dies umsetzt soll im Folgenden kurz skizziert werden.

Die forensische Sprechererkennung ist eine vergleichsweise junge Disziplin der wissenschaftlichen Kriminalistik, die sich mit ausgewählten Aspekten der Personenidentifizierung beschäftigt. In der Bundesrepublik Deutschland wird die forensische Sprechererkennung in der Regel auf Veranlassung der Landeskriminalämter (LKA) durch das Bundeskriminalamt (BKA) in Wiesbaden durchgeführt. Die Sprechererkennung kommt dabei ausschließlich bei Delikten aus dem Bereich der Schwerekriminalität zum Einsatz. Dazu zählen insbesondere Erpressung, Mord, Entführung, Terrorismus, Raub oder sexuelle Nötigung / Belästigung (vgl. Künzel 1990:1). In all diesen Fällen dient die Stimme des Täters als mittel- oder unmittelbares Tatwerkzeug. So z.B. bei mündlichen (Geld-) Forderungen bei erpresserischem Menschenraub, bei Bekennerrufen vor / nach terroristischen Akten oder bei (meist sexuell motivierten) Belästigungen.

Für den Erfolg des laufenden Ermittlungsverfahrens ist es dabei besonders günstig, wenn die Stimme des vermeintlichen Täters aufgezeichnet werden konnte. In diesen Fällen fertigt das BKA meist sehr aufwändige Stimmenanalysen (*Stimmenprofile*) an und wertet dabei zahlreiche unterschiedliche Stimmparameter aus. Liegt mindestens eine fragliche und eine bekannte Sprachaufzeichnung vor, so wird ein Stimmenvergleich durchgeführt⁸ (zum genauen Ablauf von Stimmenanalysen und Stimmenvergleichen siehe auch Heike 1990:481-489, Künzel 1987:77-102; den Fall eines Stimmenvergleichs nach der Entführung und anschließenden Ermordung eines elfjährigen Mädchens schildert Künzel 1987:5f). Problematisch ist in diesen Fällen

⁸Dabei soll geklärt werden, ob es sich um den selben oder um unterschiedliche Sprecher handelt.

5. Naive und professionelle Stimmerkennung

die häufig mangelhafte Qualität der Aufzeichnung, insbesondere bei Telefonanrufen⁹.

Aber auch ohne Stimmenaufzeichnung können beispielsweise die Opfer von Vergewaltigungen, Missbrauch oder Überfällen in manchen Fällen einen stimmlichen Eindruck des Täters erhalten. Wurde die Täterstimme also lediglich durch das Opfer oder einen Zeugen gehört, so besteht die Möglichkeit nach der Ergreifung von Verdächtigen eine so genannte Stimmenwahlgegenüberstellung (*voice-line-up* bzw. *akustische Gegenüberstellung*) durchzuführen¹⁰. Aus einer Menge vergleichbarer Sprecher, die auf Grundlage der Ergebnisse einer Stimmanalyse ausgewählt wurden, muss das Opfer / der Zeuge die „Täter-Stimme“ identifizieren (so genanntes *auditives Konfusionsexperiment*) (vgl. Künzel 1998:374, 1997:101). Dies ist mit der optischen „Täter-Opfer / Zeugen“-Gegenüberstellung vergleichbar, bei der das Opfer / der Zeuge den Täter aus einer Auswahl mehrerer Personen visuell identifizieren soll. Sowohl bei der akustischen als auch der optischen Gegenüberstellung spielt in der Kriminologie selbstverständlich auch die Glaubwürdigkeit der aussagenden Opfer / Zeugen eine nicht unerhebliche Rolle (siehe dazu z.B. Krahe & Kundrotas (1992), Steller, Wellershaus & Wolf (1992); ausführlich bei Cutler & Penrod (1990)). Damit sind weniger bewusste Falschaussagen als beispielsweise Wahrnehmungs- und Erinnerungsverzerrungen auf Grund der psychischen Stresssituation, unter der der Stimmeneindruck gewonnen wurde, gemeint. Mit diesen stressbedingten Einschränkungen und Folgen beschäftigt sich insbesondere die forensische Psychologie (Langenmayr 1997:659).

Aber auch die phonetische Textanalyse¹¹, die Qualitätsverbesserung mit digitaler Sprachsignalverarbeitung¹², die Authentisierung von Tonträgeraufzeichnungen¹³, nicht-sprachliche akustische Untersuchungen¹⁴ und die linguistische Textanalyse¹⁵

⁹In über 95 % der Fälle wird das Telefon als (Sprach- / Stimm-) Übermittler genutzt, was die Arbeit des BKA auf Grund der Beschränkung des Frequenzganges auf 300 bis 3.400 Hz, seiner Dynamikbegrenzung auf 30 dB und des additiven Rauschens nicht unwesentlich erschwert (siehe dazu u.a. auch die Untersuchungen von Künzel (2001), Cerrato, Falcone & Paoloni (2000) und Yarmey (1991) zur Stimmerkennung bei unbekanntem Anrufer). Bei Mobilfunktelefonen kommt es durch allgemein übliche Datenreduktionsverfahren (*Sprachsynthese*) zu einem noch größeren Verlust von spezifischen Sprechermerkmalen (vgl. Künzel 1988:219, 1998:370f, 2001:81ff).

¹⁰Eine solche Gegenüberstellung nach einer Vergewaltigung schildert Künzel 1997:100, siehe auch Grabe & Nolan (1996), Broeders, Cambier-Langeveld & Vermeulen (2002), Broeders & Rietveld (1995:24-40), Hollien (1990:193f).

¹¹Relevant beispielsweise bei der Auswertung aufgezeichneter Cockpitgespräche nach einem Flugzeugabsturz, wenn zur Rekonstruktion des Unfallablaufes geklärt werden muss, was zu welchem Zeitpunkt von wem gesprochen wurde.

¹²Z.B. als Vorbereitung für die o.g. Stimmenanalyse (vgl. Künzel 1997:100, Künzel 1987:7).

¹³Nachweis über Manipulationen wie nachträgliche Löschungen, Neuaufnahmen oder Schnitte (vgl. Künzel 1997:100).

¹⁴Z.B. Motorengeräusche, Schüsse etc.

¹⁵Autorenerkennung bei geschriebener Sprache; jedoch keine (Hand-) Schrifterkennung im eigentli-

5. Naive und professionelle Stimmerkennung

fallen in den Zuständigkeitsbereich der Sprechererkennung des BKA (vgl. Jessen 2004, Künzel 1997:100ff).

Die heutigen Möglichkeiten der forensischen Sprechererkennung unterscheiden sich von denen McGehees in den dreißiger und vierziger Jahren des vergangenen Jahrhunderts erheblich (siehe oben). Nach den zunächst rein auditiven Beurteilungen durch geschulte Gutachter und dem gescheiterten Versuch einer großrechnergestützten und voll automatisierten Sprechererkennung in den siebziger Jahren findet im BKA seit 1980 das *phonetisch-instrumentelle Kombinationsverfahren* Anwendung (vgl. Künzel 1995:70). Dieses Verfahren, das erfolgreich auch in andere Länder exportiert wurde, beruht auf der Kombination aus dem einschlägigen Wissen des Sprachwissenschaftlers und dem quantifizierenden akustischen Signalanalyseverfahren (vgl. Künzel 1998:381; ausführlich in Rose 2002:303-326 und Künzel 1989:29-76). Die Vorteile liegen dabei insbesondere darin, dass auf der einen Seite die quantitativen Daten Zeit sparend und exakt mit Hilfe instrumenteller und automatisierter Verfahren erhoben werden können, auf der anderen Seite hingegen insbesondere Merkmale aus dem dialektalen Bereich, dem Bereich der Sprechfehler oder der Hesitationsphänomene wesentlich zuverlässiger und effizienter durch das geschulte Ohr des Experten beurteilt werden können (Kombination objektiver und subjektiver Verfahren) (vgl. Künzel 1998:371).

Mit Hilfe dieser Untersuchungsmethode sollen identitätsrelevante Aussagen über eine sprechende Person getroffen werden. Dazu gehören insbesondere Aussagen über

- Geschlechtszugehörigkeit,
- biologisches Alter,
- regionalen oder ausländischen Akzent,
- (nicht)sprachliche (akustisch manifeste) Angewohnheiten (*habits*),
- Bildungsniveau und
- Grad der verbalen Ausdrucksfähigkeit etc. (vgl. Künzel 1998:369, 1997:100, Hollien 1990:191).

Das vorangegangene Kapitel ermöglichte zunächst einen Überblick über die durch die Wahl der Methode verursachten Beeinflussungsmöglichkeiten auf die Wiedererkennungsfähigkeit von Stimmen. Diese reichen von der Auswahl der Sprecher (z.B. berühmt, persönlich bekannt, unbekannt), über die verwendeten Stimuli (Dauer, Repertoire) und Methoden (Identifizieren, Erkennen, Vergleichen) bis zur Auswahl des

chen Sinne.

5. Naive und professionelle Stimmerkennung

Hörerensembles (Kinder oder Erwachsene mit / ohne Beeinträchtigung). In einem weiteren Abschnitt wurden mehrere Studien vorgestellt, die eine Identifizierung der für das Stimmerkennen relevanten akustischen Parameter zum Ziel hatten. Dabei kristallisierte sich heraus, dass die holistische Gestalt einer Stimme anhand unterschiedlicher und von Sprecher zu Sprecher möglicherweise variierender Parameter erfolgt. Zu den wesentlichen Faktoren gehören neben der mittleren Sprechstimmlage, auch die Formantfrequenzen und die Transitionsbewegungen der Vokoide. Die für die Erkennungsleistung rechts- bzw. linkshirngeschädigter Patienten relevanten Untersuchungen werden in Kapitel 7 vorgestellt.

6. Anatomie und Physiologie der peripheren und zentralen Hörorgane

6. Anatomie und Physiologie der peripheren und zentralen Hörorgane

Nachdem der Sprechschall produziert und durch das Medium Luft weiter getragen wird, erreicht er das Ohr des Hörers. Mit der letzten Stufe der Kommunikationskette beschäftigt sich das folgende Kapitel, das einen Überblick über die anatomisch-physiologischen Grundlagen des Hörvorganges gibt. Während man unter dem Begriff *Ohr* im Allgemeinen lediglich einen Schallaufnahme- und -umsetzungsapparat versteht, der den Sprechschall in elektrische Impulse umwandelt, werden unter dem Begriff *Hörorgan* auch die für den Hörvorgang erforderlichen Anteile des Zentralnervensystems verstanden. Man unterscheidet demnach in peripheres und zentrales Hörorgan (vgl. Pétursson & Neppert ³2002:170).

6.1. Die peripheren Anteile des Ohres

Die peripheren Anteile des Ohres bestehen aus dem Außen-, Mittel- und Innenohr.

6.1.1. Das Außenohr (*Auris externa*)

Das Außenohr setzt sich aus der Ohrmuschel (*Auricula*) und dem äußeren Gehörgang (*Meatus acusticus externus*) zusammen. Die Ohrmuschel, die aus einem elastischen Knorpelgerüst besteht und u.a. den äußeren Gehörgang schützen soll, sorgt durch Reflektion der Schallwellen für eine geringfügige aber messbare Erhöhung der Schallintensität und spielt eine wesentliche Rolle für das Richtungs- und Raumhören. Der sich anschließende äußere Gehörgang ist mit Epidermis ausgekleidet und umfasst eine Länge von etwa 2,5 cm. Er führt in einer leichten Krümmung zu dem schräg eingelassenen Trommelfell (*Membrana tympanica*), das Außen- und Mittelohr voneinander trennt (vgl. Trepel ³2004:313f, Pétursson & Neppert ³2002:170ff, Faller & Schünke ¹⁴2004:718, Neppert ⁴1999:274f).

6.1.2. Das Mittelohr (*Auris media*)

Das Mittelohr beginnt am Trommelfell und besteht in erster Linie aus der Paukenhöhle (*Cavitas tympani*). Sie ist ein luftgefüllter Hohlraum der eine Kette aus drei Gehörknöchelchen beinhaltet. Diese Kette setzt sich aus dem Hammer (*Malleus*), dem Amboss (*Incus*) und dem Steigbügel (*Stapes*) zusammen (vgl. Abbildung 6.2). Sie bilden gemeinsam mit dem Trommelfell, dem äußeren Gehörgang und der Ohrmuschel den Schalleitungsapparat. Über die Schwingungen des Trommelfells und das sich anschließende Hebelsystem der Gehörknöchelchen werden die Schallwellen vom gasförmigen Medium Luft auf das flüssige Medium Perilymphe im Innenohr weitergeleitet. Da das Trommelfell mit einer (schwingenden) Fläche von etwa 0,55 cm² rund zwanzig mal größer ist als die Fläche des ovalen Fensters (*Fenestra cochlea*) (ca.

6. Anatomie und Physiologie der peripheren und zentralen Hörorgane

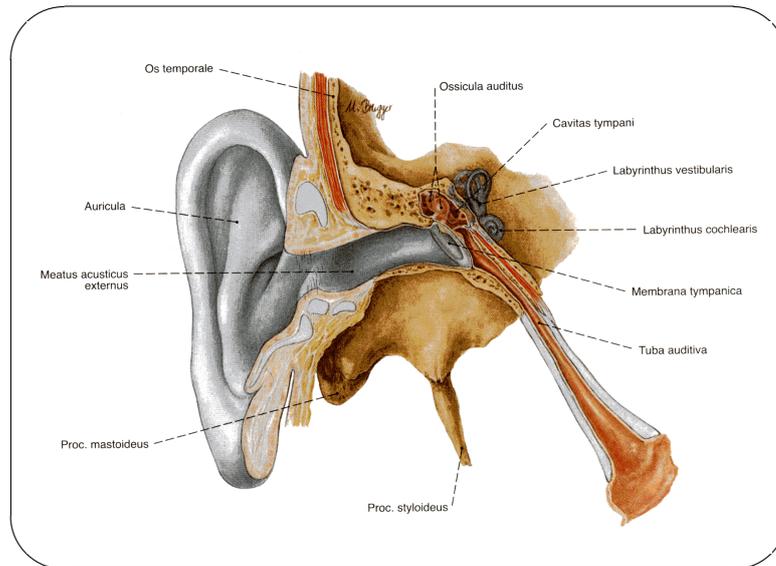


Abbildung 6.1.: Ohr (halbschematisch), teilweise eröffnet von frontal (rechts)
(aus: Putz & Pabst ²¹2004:381)

0,03 cm²) wird eine Schwingungswechseldruckverstärkung von etwa 25 dB erzielt. Auch die Hebellängenverhältnisse der Knöchelchen bewirken eine leichte Erhöhung des Verstärkungsfaktors, so dass es insgesamt zu einer Verstärkung des Schwingungsdrucks von etwa 26 bis 27 dB kommt. Die Ohrtrumpete (*Tuba auditiva*) verbindet die Paukenhöhle mit dem Nasenschlund (*Nasopharynx*), sorgt u.a. für eine Belüftung des Mittelohrs und ermöglicht einen Druckausgleich zur äußeren Atmosphäre (vgl. Trepel ³2004:315-318, Pétursson & Neppert ³2002:172, Faller & Schünke ¹⁴2004:718f, Neppert ⁴1999:276f).

6.1.3. Das Innenohr (*Auris interna*)

Das Innenohr setzt sich aus dem Vorhof (*Vestibulum*) und dem knöchernen Labyrinth (*Labyrinthus osseus*) zusammen. In letzteres ist das mit Endolymphe gefüllte häutige Labyrinth (*Labyrinthus membranaceus*) eingelagert, das das Gleichgewichts- und Gehörorgan enthält: die Bogengänge (*Labyrinthus vestibularis*) und die Schnecke (*Cochlea*) (vgl. Abbildung 6.3 auf Seite 75). Sie wandelt die mechanischen Schwingungen in Nervensignale um und ermöglicht eine erste Schallanalyse. Die (*Cochlea*) hat die Form einer Schnecke mit $2\frac{1}{2}$ Windungen um eine knöcherne Achse (*Modiolus*) und wird in ganzer Länge in ihrer Mitte von der *Reissnerschen Membran* und der *Basilarmembran*, die gemeinsam den Endolymphschlauch *Ductus cochlearis* bilden, in zwei Kanäle geteilt. An der Spitze sind beide Kanäle (oder Treppen) miteinander verbunden (*Helicotrema*). Der obere Kanal wird als Vorhoftreppe (*Scala vestibuli*),

6. Anatomie und Physiologie der peripheren und zentralen Hörgane

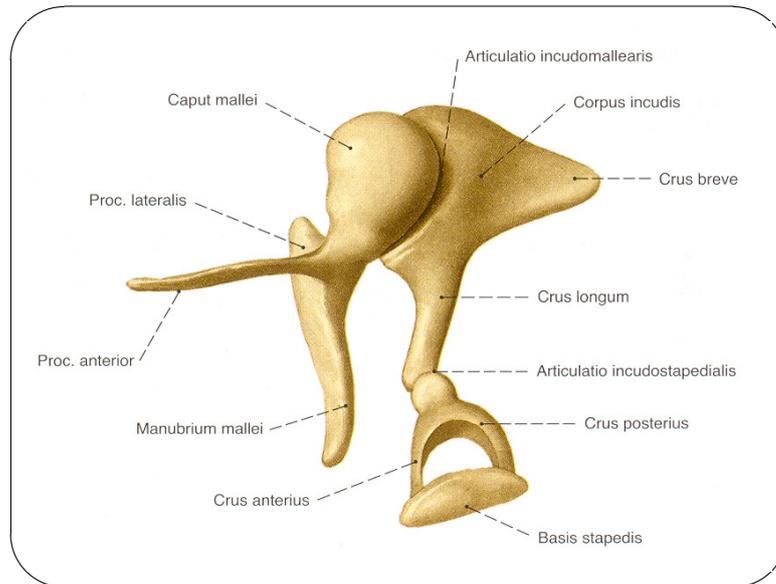


Abbildung 6.2.: Gehörknöchelchen, von oben medial (rechts)
(aus: Putz & Pabst ²¹2004:386)

der untere als Paukentreppe (*Scala tympani*) bezeichnet. Dieser Schlauch ist mit Endolymphe gefüllt, die durch Verschieben die so genannten Haarzellen stimuliert. Haar- und Stützzellen bilden auf der Basilarmembran das eigentliche Hörorgan, das *Cortische* oder *Corti-Organ* (vgl. Trepel ³2004:318, Faller & Schünke ¹⁴2004:719f). Bei den Haarzellen handelt es sich um Nervenendzellen (Sensoren), die bei Stimulierung in den sich anschließenden Neuronen Aktionspotenziale auslösen, da sie mit den Zellkörpern der Neurone im *Ganglion cochleare* (*Ganglion spirale cochlea*) in synaptischem Kontakt stehen. Dies geschieht, wenn die von außen ankommenden Schallwellen über die Bewegungen des Steigbügels Druckschwankungen erzeugen und den Endolymph-Schlauch in Schwingungen versetzt. Diese primäre Auslenkung an der Basis setzt sich als Wanderwelle fort und läuft, je nach Frequenz, unterschiedlich weit in die Spitze der Schnecke hinein. Bei den niedrigsten Frequenzen wandert die Welle über die gesamte Länge der Basilarmembran, bei den höchsten Frequenzen schwingt nur der schmalste Teil an der Basis. Diese Erklärung der Schallanalyse wird als *Frequenz-Ortsabbildung*, *Tonotopie* oder *Einortstheorie* bezeichnet (vgl. Neppert ⁴1999:282, Pompino-Marschall ²2003:147ff, Faller & Schünke ¹⁴2004:722). Dieses Phänomen beschrieb als erster der Mediziner und Naturwissenschaftler Hermann Ludwig Ferdinand von HELMHOLTZ (1821-1894) in seiner *Resonanzhypothese* von 1862.

Die entstandenen chemo-elektrischen Vorgänge (Aktionspotenziale) werden vom Hörnerv (*Nervus cochlearis*) zum Gehirn weitergeleitet. Gemeinsam mit dem *Nervus*

6. Anatomie und Physiologie der peripheren und zentralen Hörgane

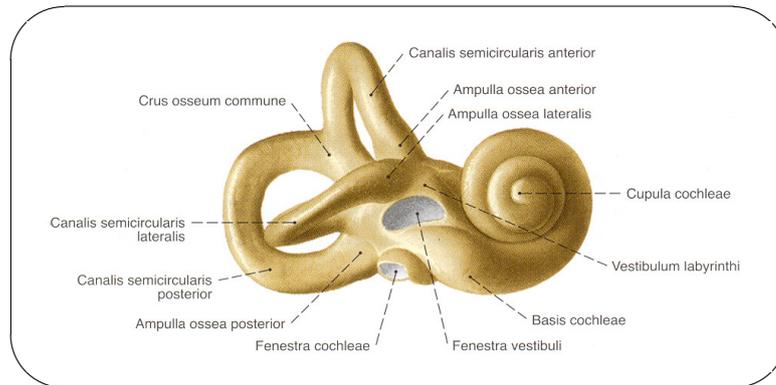


Abbildung 6.3.: Knöchernes Labyrinth mit Cochlea)
(aus: Putz & Pabst ²¹2004:393)

vestibularis, der im vestibulären Teil des Innenohrs entspringt, bilden sie den **VIII. Hirnnerv** (*Nervus vestibulocochlearis*). Im knöchernen inneren Gehörgang des Felsenbeins treten sie zu einem Nervenstamm zusammen und ziehen durch diesen bis zum Hirnstamm (*Truncus cerebri*), der sich aus dem verlängerten Mark (*Medulla oblongata*), der Brücke (*Pons*) und dem Mittelhirn (*Mesencephalon*) zusammensetzt. Hier trennen sie sich wieder und enden im Pons in den entsprechenden Hirnnervenkernen, den *Nuclei cochleares* bzw. den *Nuclei vestibulares*. Ausgehend vom *Nuclei cochleares dorsalis* führen die meisten Axone nach kontralateral und steigen von dort zur Vierhügelplatte (*Tectum*) zum Kern des *Colliculus inferior* (*Nucleus colliculi inferioris*) auf. In dieses starke Faserbündel (*Corpus trapezoideum*) sind weitere Kerne eingeschaltet, die u.a. für das Richtungshören von Bedeutung sind (*Ncl. olivaris superior*). Andere Fasern bleiben ipsilateral und enden dann ebenfalls im Nucleus colliculi inferioris. Gemeinsam bilden sie die zentrale Hörbahn (*Lemniscus lateralis*). Nach weiteren Verschaltungen verlaufen die Hörbahnfasern durch den hinteren Abschnitt der *Capsula interna* als Hörstrahlung zur primären Hörrinde (vgl. Trepel ³2004:65, Gertz ³2001:94). Bei jeder Verschaltung behält die Hörbahn ihre topotopische Gliederung bei, was bedeutet, dass jede Tonfrequenz einen eigenen Terminationsort in der primären Hörrinde zeigt (vgl. Trepel ³2004:322, Neppert ⁴1999:287). Tiefe Frequenzen rostro-lateral, höhere Frequenzen kaudo-medial (vgl. Bähr & Frotscher ⁸2003). Dies bildet die Grundlage für der Tonhöhenwahrnehmung (vgl. Kapitel 3.2.2).

6.2. Primäre und sekundäre Hörrinde

Die **primäre Hörrinde** entspricht dem Brodmann-Areal 41 und besteht aus den Heschl'schen-Querwindungen (*Gyri temporalis transversi*, vgl. Abbildung 6.4), die, wie der Name nahe legt, quer zu allen anderen Gyri verlaufen. Sie liegen unterhalb des parietalen und frontalen Operculums auf der dorsalen Seite des Temporallappens (*Lobus temporalis*). Über die primäre Hörrinde werden zunächst nur einzelne Laute oder Lautmuster unterschiedlicher Frequenzen wahrgenommen, jedoch keine Wörter oder Melodien. Sie dient lediglich der interpretationsfreien Bewusstwerdung auditorischer Impulse aus dem Innenohr. Eine sinnvolle Verknüpfung einzelner Laute zu Wörtern oder Sätzen erfolgt dann erst in der sekundären Hörrinde. Sie ist das Ziel der efferenten Bahnen der primären Hörrinde.

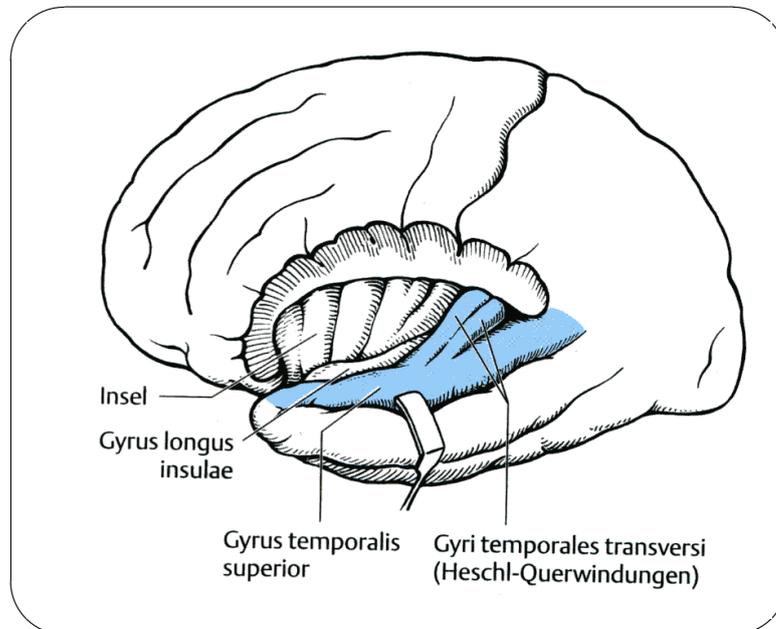


Abbildung 6.4.: Heschl'sche Querwindung (*Gyrus temporalis transversi*)
(aus: Bähr & Frotscher ⁸2003:358)

Die **sekundäre Hörrinde** erstreckt sich über die Brodmann-Areale 42 und 22 und grenzt somit an die Heschl'schen-Querwindungen bzw. die primäre Hörrinde. Wie bereits an anderer Stelle erwähnt, erhält sie von hier auch den Großteil ihrer Afferenzen. Haben im Laufe der Entwicklung entsprechende Lernprozesse stattgefunden, so erfahren die auditorischen Impulse hier eine interpretative Verarbeitung: z.B. werden Laute als Wörter und Tonfolgen als Melodien erkannt. Dieser Leistung

6. Anatomie und Physiologie der peripheren und zentralen Hörorgane

zugrunde liegen anatomische Schaltkreise, die ein Zuordnen des Gehörten zu bereits Gelerntem ermöglichen. Die primäre, aber vor allem auch die sekundäre Hörrinde hat intensive Afferenzen mit dem *Gyrus angularis*, der am hinteren Ende des *Sulcus lateralis* des Temporallappens die Area 39 umfasst. Dieser spielt eine wesentliche Rolle bei der Verknüpfung von visuellem und auditorischem Input und gilt als Schaltstelle zwischen visuellem und sprachlichem Kortex. Eine Läsion im *Gyrus angularis* resultiert klinisch z.B. in Störungen des Lesens und / oder des Schreibens (*Alexie*, *Agraphie*). Darüber hinaus zeigt die sekundäre Hörrinde der sprachdominanten Hemisphäre auch efferente Verbindungen zu zahlreichen kortikalen Assoziationsfeldern. Über den *Fasciculus arcuatus* (*Fibrae arcuatae cerebri*) verlaufen Verbindungen zum motorischen Sprachzentrum (Broca-Areal) (vgl. Trepel ³2004:230).

Dieses Kapitel erläuterte überblicksweise die (neuro-) anatomischen Grundlagen des peripheren und zentralen Hörorgans. Für den weiteren Verlauf der Arbeit sind insbesondere die kortikalen Verarbeitungsprozesse akustischer Impulse von herausragender Bedeutung.

7. Neurolinguistische und -psychologische Grundlagen

7. Neurolinguistische und -psychologische Grundlagen

Die Neurolinguistik ist eine stark interdisziplinär ausgerichtete Teildisziplin der Linguistik und integriert u.a die relevanten Erkenntnisse der Neurologie, der Neurobiologie und der Neuropsychologie. Sie beschäftigt sich u.a mit den neurologischen Grundlagen der Sprache und den zerebral bedingten Sprach- und Sprechstörungen. (vgl. Müller 2002:406, Peuser & Winter 2000:133).

Die Neuropsychologie ist eine Teildisziplin der Klinischen Psychologie und beruht auf Ansätzen der Allgemeinen, der Experimentellen und der Physiologischen Psychologie. Sie beschäftigt sich mit den so genannten *höheren Hirnleistungen*, worunter man u.a. Sprache, Wahrnehmung, Kognition, Emotion und Bewusstheit versteht. Ihre Anfänge gehen auf den deutschen Arzt und Anatom Franz Joseph GALL (1758-1828) zurück, der die *Gall'sche Schädellehre* bzw. *Phrenologie* begründete, die aus der Form des Schädels u.a. auf geistige und seelische Anlagen zu schließen versuchte.

Ein Teilgebiet der Neuropsychologie ist die die **Klinische Neuropsychologie**, die sich mit den Folgen zentralnervöser Schädigungen auf das menschliche Verhalten und Empfinden beschäftigt. Die durch experimentelle Forschung gewonnenen Erkenntnisse werden für die Diagnostik und Therapie von Patienten mit Hirnfunktionsstörungen eingesetzt.

Nach einer Einführung in die Phänomenologie der aphasischen Standard- und Sonderformen als Folgen linkshemisphärischer Läsionen folgt eine Darstellung der Störungen der Objekterkennung. Das Kapitel schließt mit einer Betrachtung der auditiven Agnosien, wobei der Schwerpunkt speziell auf der für die vorliegende Arbeit relevanten Phonagnosie liegt.

7.1. Grundlagen der Aphasie

Der Begriff *Aphasie* bedeutet übersetzt *Sprachlosigkeit* und wurde von dem französischen Neurologen Armand TROUSSEAU eingeführt. Sie beschreibt einen Symptomenkomplex, der

- als Folge einer plötzlich auftretenden kortikalen Läsion der sprachdominanten Hemisphäre eintritt,
- sprachliche Leistungen betrifft, während Gedächtnis und Intelligenz relativ intakt bleiben,
- nach abgeschlossenem Spracherwerb auftritt,
- gewöhnlich alle sprachlichen Modalitäten und Ebenen betrifft und
- relativ stabil und chronisch verläuft (nach Tesak ²2006:2).

7. Neurolinguistische und -psychologische Grundlagen

In etwa 65 bis 80 % der Aphasien liegt die Ursache in apoplektischen Insulten („Schlaganfälle“). Auslöser dafür sind i.d.R. linkshemisphärische ischämische Hirninfarkte, Hirnblutungen oder Subarachnoidalblutungen. Unter Hirninfarkten versteht man zerebrale Durchblutungsstörungen, die zu einer Nekrose des Gewebes führen. Die Ursachen solcher Durchblutungsstörungen liegen entweder in Thrombosen, also lokalen Gerinnselbildungen, oder Embolien, die durch verschleppte Gerinnsel entstehen können. Weitere Auslöser für Aphasien können Schädel-Hirn-Traumata, Hirntumoren, Atrophien, oder Enzephalitiden sein (vgl. Leischner ²1987:68f). Diese können jedoch zu einer Vermischung der Syndrome und verstärkt zu nicht-sprachlichen Symptomen führen (vgl. Huber, Poeck & Weniger ⁵2002:121).

Der neoklassischen Lehrmeinung und dem Aachener Klassifikationsschema folgend, werden im nächsten Unterkapitel die aphasischen Standard- und Sonderformen vorgestellt. Dieser neoklassische Ansatz basiert auf den Annahmen von Franz Joseph GALL (1758-1828), Pierre Paul BROCA (1824-1880), Carl WERNICKE (1848-1905) und Ludwig LICHTHEIM (1845-1928), die annahmen, dass je nach Läsionsort im Gehirn des Patienten spezifische Symptomenbündel zu beobachten seien (vgl. Tesak ²2006:26).

7.1.1. Standardsyndrome

Zu den so genannten Standardsyndromen gehören die:

- amnestische Aphasie,
- Broca-Aphasie,
- Wernicke-Aphasie und
- globale Aphasie.

Amnestische Aphasie

Die Sprache der Betroffenen zeigt i.d.R. einen relativ intakten Satzbau und zahlreiche Wortfindungsstörungen (insbesondere beim Benennen). In den meisten Fällen gelingt es den Patienten jedoch, mit Hilfe von Umweg- und Ersatzstrategien diese Defizite zu kompensieren. Zusätzlich zeigen sie Satzabbrüche, semantische Paraphasien mit geringer Abweichung vom Zielwort sowie vereinzelt auch phonematische Paraphasien bei Inhaltswörtern. Da das Sprachverständnis der Patienten meist nur wenig beeinträchtigt ist, zeigen sich nur geringfügige Einschränkungen bei der Kommunikationsfähigkeit (vgl. Tesak ²2006:28, Huber, Poeck & Weniger ⁵2002:121, Schöler & Grötzbach 2002:26, Biniek ²1997:2).

7. Neurolinguistische und -psychologische Grundlagen

Der Symptomenkomplex der amnestischen Aphasie lässt sich nur schwer der Läsion einer bestimmten Hirnregion zuordnen. Meist handelt es sich um kleinere temporoparietale Läsionen im kortikalen oder subkortikalen Bereich (vgl. Huber, Poeck & Weniger ⁵2002:121). Auch Tesak berichtet von retrorolandisch lokalisierten Läsionen im hinteren Teil des Temporal- oder unteren Bereich des Parietallappens sowie von Läsionen im Bereich des Gyrus angularis (²2006:38) (siehe auch Abbildung 7.1 auf Seite 84).

Broca-Aphasie

Die Spontansprache der Broca-Aphasiker ist stockend, angestrengt und mitunter stark dysprosodisch. Häufig kommt es zu phonematischen Paraphasien, agrammatischem Satzbau und reduzierten Satzkonstruktionen (Ein- / Zweiwortsätze mit syntaktischen Mängeln). Bei schriftsprachlichen Äußerungen zeigen Broca-Aphasiker häufig graphematische Paraphasien und Agrammatismus. Phonationsprobleme, undeutliche Artikulation und eine Nivellierung der Sprechmelodie können auf eine bestehende Dysarthrie hinweisen, die in manchen Fällen begleitend auftreten kann. Die mündliche Kommunikation variiert mit dem Schweregrad der expressiven Leistung, ist jedoch, v.a. auf Grund des relativ gut erhaltenen Sprachverständnisses, meist möglich. Als problematisch erweisen sich mitunter jedoch die dysarthrischen Verschleifungen und Entstellungen der Äußerungen (vgl. Tesak ²2006:28, Huber, Poeck & Weniger ⁵2002:125, Schöler & Grötzbach 2002:26, Biniek ²1997:3).

Die der Broca-Aphasie typischerweise zugeordneten, relativ großen Läsionsorte liegen im Versorgungsgebiet der *Arteria prearolandica* bzw. der vorderen Äste der mittleren Hirnschlagader. Sie umfassen den suprasylvischen, prärolandischen Bereich, unter einer möglichen Beteiligung des Broca-Areals (Brodmann-Area 44 und 45). Die Computertomogramme betroffener Patienten zeigen i.d.R. Läsionen dorsal der von BROCA ursprünglich beschriebenen Region im Marklager des Stirnhirns, die auch die Insula betreffen können (vgl. Huber, Poeck & Weniger ⁵2002:124, Tesak ²2006:38) (siehe auch Abbildung 7.1 auf Seite 84).

Wernicke-Aphasie

Spontansprachlich zeigen Wernicke-Aphasiker eine überschießende, logorrhoeische Sprachproduktion mit häufigen phonematischen und semantischen Paraphasien und Neologismen, die bis zum phonematischen bzw. semantischen Jargon führen können. Die Satzebene ist geprägt von paragrammatischen Erscheinungen. Das Sprachverständnis, insbesondere im Bereich der Inhaltswörter, ist stark eingeschränkt. Dies

7. Neurolinguistische und -psychologische Grundlagen

gilt auch für den Bereich der Schriftsprache. Das Kommunikationsverhalten der Wernicke-Aphasiker erscheint auf den ersten Blick vermeintlich intakt, ist jedoch bei genauer Prüfung stark eingeschränkt (vgl. Tesak ²2006:28, Huber, Poeck & Weniger ⁵2002:1, Schöler & Grötzbach 2002:26, Biniek ²1997:4).

Wernicke-Aphasiker zeigen i.d.R. eine Läsion im Versorgungsbereich der *Arteria temporalis posterior* aus der *Arteria cerebri media*. Dabei ist überwiegend der hintere Teil des *Gyrus temporalis superior* betroffen. Zeigt der Patient vorwiegend phonematische Paraphasien, so ist die Läsion eher parietal lokalisiert; überwiegen die semantischen Paraphasien, so erstreckt sie sich eher nach temporal (Vignolo 1988, nach Huber, Poeck & Weniger ⁵2002:132) (siehe auch Abbildung 7.1 auf Seite 84).

Globale Aphasie

Die globale Aphasie gilt als die schwerste Form der Aphasien. Die Kommunikationsmöglichkeiten der Betroffenen sind stark eingeschränkt, da häufig nur Einzelwörter und Sprachautomatismen aktiviert werden können, die meist zusätzlich perseveriert werden. Die Modalitäten Sprachverständnis, Lesen und Schreiben gelten als stark beeinträchtigt, so dass eine Kommunikation oftmals nur sehr schwer möglich ist (vgl. Tesak ²2006:28, Huber, Poeck & Weniger ⁵2002:139, Schöler & Grötzbach 2002:26, Biniek ²1997:4).

Meistens wird die globale Aphasie durch einen embolischen oder thrombotischen Verschluss des Hauptstammes der *Arteria cerebri media* verursacht und betrifft somit die komplette perisylvische Region. Die Läsionen können dabei auch umfassend in die weiße Substanz reichen (vgl. Tesak ²2006:38, Huber, Poeck & Weniger ⁵2002:132) (siehe auch Abbildung 7.1 auf Seite 84).

7.1.2. Sonderformen

Zu den Sonder- bzw. Nichtstandardsyndromen zählen die Bostoner und Aachener Schule vier weitere Symptomenbündel. Anders als bei den o.g. Standardsyndromen spielen im Rahmen der Diagnostik einzelne Leistungen – insbesondere die Leistung in der Modalität Nachsprechen – eine herausragende Rolle.

Leitungsaphasie

Patienten mit einer Leitungsaphasie zeigen spontansprachlich zahlreiche phonematische Paraphasien. Die übrigen Symptome ähneln denen der Wernicke-Aphasie, das Sprachverständnis ist im Vergleich jedoch weniger beeinträchtigt. Besonders auffällig ist die stark eingeschränkte Fähigkeit, vorgegebene Sätze nachzusprechen. Dabei neh-

7. Neurolinguistische und -psychologische Grundlagen

men die phonematischen Entstellungen mit der Länge des nachzusprechenden Materials zu (vgl. Tesak ²2006:29, Huber, Poeck & Weniger ⁵2002:147, Biniak ²1997:5).

Bei der Leitungsaplasie sind jene Regionen betroffen, die das Broca- und Wernicke-Areal miteinander verbinden, also die weiße Substanz unter dem *Gyrus supramarginalis* (entspricht dem *Fasciculus arcuatus*) (vgl. Tesak ²2006:38, Huber, Poeck & Weniger ⁵2002:146) (siehe auch Abbildung 7.1 auf der nächsten Seite).

Transkortikal-motorische Aphasie

Die Spontansprache der Betroffenen ist sehr stark eingeschränkt, so dass sie sich kaum oder gar nicht spontan äußern. Im auffälligen Gegensatz dazu können sie jedoch mit intakter Syntax und deutlicher Artikulation vorgegebenes Satzmaterial nachsprechen. Das Sprachverständnis ist gut erhalten und die Patienten sind in der Lage laut vorzulesen (vgl. Tesak ²2006:29, Huber, Poeck & Weniger ⁵2002:147, Biniak ²1997:5f).

Die Läsionen bei der transkortikal-motorischen Aphasie liegen typischerweise im supplementären motorischen Kortex bzw. im Frontallappen anterior des Broca-Areals (vgl. Tesak ²2006:38, Huber, Poeck & Weniger ⁵2002:147) (siehe auch Abbildung 7.1 auf der nächsten Seite).

Transkortikal-sensorische Aphasie

Die Symptomatologie der Patienten ähnelt der der Wernicke-Aphasie. Spontansprachlich zeigen sie zahlreiche semantische Paraphasien und Echolalien. Ihre Leistungen im Bereich Nachsprechen sind gut erhalten, jedoch kann i.d.R. der Sinn des Nachgesprochenen nicht oder nur stark eingeschränkt erfasst werden (vgl. Tesak ²2006:29).

Bei transkortikal-sensorischen Patienten finden sich meist Läsionen im temporo-okzipitalen Marklager ohne Beteiligung der Wernicke-Region. Es handelt sich dabei um das Grenzgebiet der *Arteria cerebri media* und der *Arteria cerebri posterior*. Vereinzelt wird auch über transkortikal-sensorische Aphasiker mit Läsionen frontaler Hirnareale berichtet (vgl. Tesak ²2006:38, Huber, Poeck & Weniger ⁵2002:148) (siehe auch Abbildung 7.1 auf der nächsten Seite).

Gemischte transkortikale Aphasie

Patienten mit einer gemischten transkortikalen Aphasie können meist relativ gut nachsprechen, während die Spontansprache angestrengt, stockend und häufig echolalisch ist und zahlreiche Automatismen und Stereotypen aufweist. Zusätzlich ist

7. Neurolinguistische und -psychologische Grundlagen

das Sprachverständnis deutlich eingeschränkt (vgl. Tesak ²2006:29, Huber, Poeck & Weniger ⁵2002:148).

Die Läsionen der gemischten transkortikalen Aphasie betreffen i.d.R. die Verbindungen zwischen der perisylvisch lokalisierten Sprachregion und den entsprechenden Assoziationskortexen. In den letzten Jahren wurde jedoch wiederholt von Patienten berichtet, die nach ausgedehnten Mediainfarkten die Symptome einer gemischten transkortikalen Aphasie zeigten, was die neoklassische Annahme nicht erklären kann (vgl. Huber, Poeck & Weniger ⁵2002:148, siehe auch Abbildung 7.1).

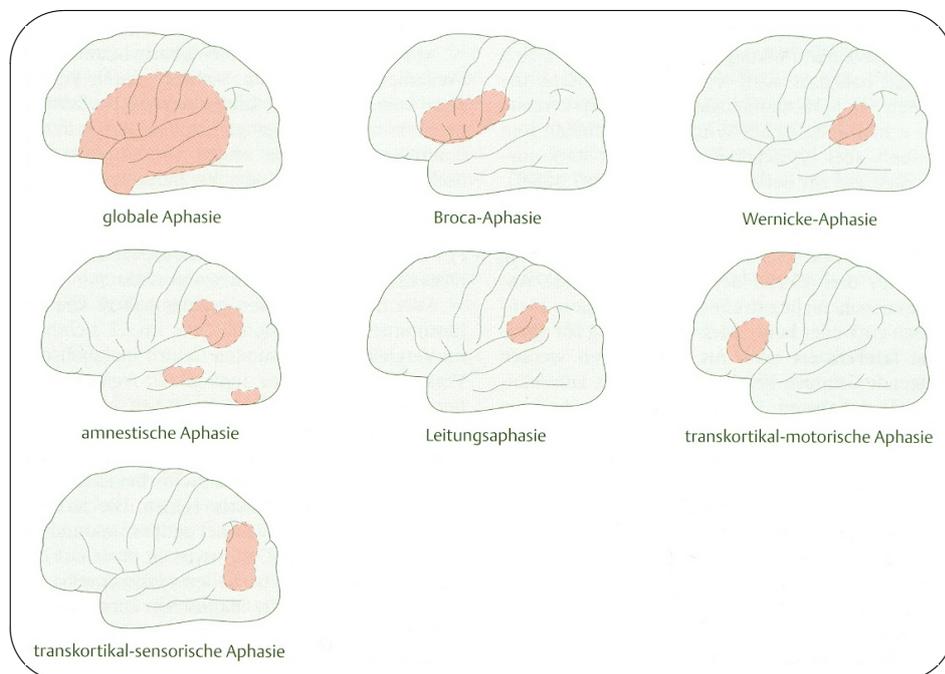


Abbildung 7.1.: Läsionsorte bei aphasischen Syndromen
(aus: Tesak ²2006:39)

7.2. Störungen der Objekterkennung (*Agnosien*)

Störungen der Objekterkennung werden als *Agnosien* bezeichnet. In Anlehnung an Frederiks (1969) formulieren Bauer & Demery (⁴2003):

„*Agnosia* is a relatively rare neuropsychological symptom defined in the classical literature as a failure of recognition that cannot be attributed to elementary sensory defects, mental deterioration, attentional disturban-

7. Neurolinguistische und -psychologische Grundlagen

ces, aphasic misnaming, or unfamiliarity with external stimuli“ (Bauer & Demery 2003:236, Hervorhebung im Original).

Agnosien treten meist modalitätsspezifisch auf, wobei der Patient die Objekterkennung evtl. über einen anderen Wahrnehmungskanal kompensieren kann. So wäre beispielsweise ein Patient mit einer visuellen Agnosie möglicherweise zwar nicht in der Lage einen Würfel durch bloßes Ansehen zu erkennen. Nähme er ihn jedoch in die Hand, würde er ihn ggf. als einen solchen identifizieren können. Die häufigsten und am besten erforschten Agnosien sind die visuellen, die auditiven und die taktilen Agnosien. In der Literatur finden sich jedoch auch Studien zur Anosognosie (vgl. u.a. Cosentino & Stern 2005, Kashiwa et al. 2005) oder zu Sonderformen wie der gustatorischen Agnosie (vgl. Small et al. 2005).

7.3. Auditive Verarbeitungsstörungen

Die Ursachen für auditive Verarbeitungsstörungen können sehr unterschiedlich sein. Eine grundsätzliche Differenzierung erfolgt durch die Einteilung in periphere und zentrale Störungen der Verarbeitung. Während periphere Störungen beispielsweise durch Verletzungen und Schädigungen des Mittel- und Innenohres verursacht werden können (vgl. Kapitel 6.1.2 bzw. 6.1.3), liegt die Ursache zentral bedingter Verarbeitungsstörungen in einer Läsion rostral des *Nucleus cochlearis*, meist innerhalb des Temporallappens und / oder der perisylvischen Region (siehe unten). Schädigungen des Mittel- und Innenohres sind dem Fachgebiet der Hals-Nasen-Ohrenheilkunde und ggf. der Pädaudiologie zuzuordnen und sollen an dieser Stelle nicht weiter erläutert werden. Die häufigste Form der zentral-auditiven Verarbeitungsstörungen ist die *auditive Agnosie*. Sie ist Gegenstand des folgenden Unterkapitels.

7.3.1. Auditive Agnosien

Der Begriff *Agnosie* stammt von dem griechischen *agnōsia* (= Unkenntnis) und setzt sich aus der Vorsilbe A- (= Un-, -los, -leer) und dem Wortteil -gnos (= Kenntnis) zusammen. Er beschreibt eine Störung jener Verarbeitungsstufen bzw. -prozesse, die die akustischen Parameter eines spezifischen Signals verarbeiten. Dazu gehört z.B. der temporäre Verlauf, die spektrale Zusammensetzung oder die Kombination mehrerer Parameter zu einem Muster. Trotz erhaltener Hörfähigkeit und einer befundfreien Tonaudiometrie¹ sind die Patienten nicht in der Lage, komplexe akustische Signale zu dekodieren und ihnen einen semantischen Inhalt zuzuordnen. Dies kann sowohl

¹Mit Ausnahme bei der *kortikalen Taubheit* (vgl. Seite 86).

7. Neurolinguistische und -psychologische Grundlagen

verbales als auch nonverbales Geräuschmaterial betreffen. Da eine standardmäßige Überprüfung dieser Leistungen im Rahmen neurologischer bzw. neuropsychologischer Screenings nicht vorkommt, gibt es keine gesicherten Angaben zur Häufigkeit der Störung (vgl. Engelen et al. 1995:1396); Beschreibungen des Symptoms beschränken sich häufig auf Einzelfallstudien (vgl. u.a. Uttner et al. 2006, Kaga et al. 2005, Hattiangadi et al. 2005, Pan, Kuo & Hsieh 2004, Szirmai, Farsang & Csuri 2003, Engelen et al. 1995). In den meisten Fällen werden auditive Agnosien durch bilaterale ischämische oder hämorrhagische Insulte der Temporallappen verursacht, die zu Läsionen der Hörkortizes bzw. der subkortikalen Anteile der Hörbahn führen (vgl. u.a. Engelen 2003:141, Kaga et al. 2005:1351, Kazui et al. 1990:476, Motomura et al. 1986:379). In seltenen Fällen können auch Enzephalitiden (vgl. z.B. Kaga et al. 2003:232, Orimo et al. 2000:256), Schädel-Hirn-Traumata oder Medikamentenintoxikationen zu einer auditiven Agnosie führen (vgl. Engelen ⁵2002:436). Eine detaillierte Übersicht über die historische Entwicklung und die – meist uneinheitliche Verwendung – des Begriffs finden sich bei Vignolo (1969:172-208).

Kortikale Taubheit

Im Gegensatz zu den übrigen auditiven Agnosien haben die Patienten mit kortikaler Taubheit kein befundfreies Tonaudiogramm. Die tonaudiometrischen Schwellen sind entweder nicht messbar oder stark erhöht. Dabei ist ungeklärt, ob die Betroffenen tatsächlich das Signal wahrnehmen oder lediglich die es begleitenden Vibrationen spüren (vgl. Engelen ⁵2002:439). Die kortikale Taubheit tritt ausschließlich nach beidseitiger Läsion der auditorischen Kortizes bzw. der *Radiatio acustica* auf. Die betroffenen Patienten zeigen keinerlei Reaktion auf akustische Stimuli, auch keine Startle- und / oder Orientierungsreaktion. Sie sind weder in der Lage unterschiedliche Geräusche oder Sprache zu diskriminieren (z.B. im Rahmen einer gleich / ungleich-Entscheidung), noch zu identifizieren (vgl. Szirmai, Farsang & Csuri 2003, Engelen ⁵2002:439, Taniwaki et al. 2000). Coslett, Brashear & Heilman (1984) zeigten, dass – im Gegensatz zu den anderen auditiven Agnosien (vgl. z.B. Michel, Peronnet & Schott 1980, Jerger et al. 1969) – bei Patienten mit kortikaler Taubheit die akustisch evozierten Hirnstammpotenziale normal, die kortikalen Potenziale im EEG hingegen nicht nachweisbar sind. Dies erlaubt auch eine differenzialdiagnostische Unterscheidung zwischen der kortikalen Taubheit und der peripheren Schwerhörigkeit bzw. Ertaubung (vgl. Engelen ⁵2002:439).

Eine Sonderform bildet das *taube Hören*. Hier können Patienten durch das willentliche Fokussieren der Aufmerksamkeit eine rudimentäre Hörfähigkeit erreichen

7. Neurolinguistische und -psychologische Grundlagen

(vgl. Engelien et al. 2000). Diese Patienten sind jedoch auch im Rahmen von 2-AFC-Aufgaben nicht in der Lage, komplexe Muster oder isolierte akustische Parameter zu unterscheiden (vgl. Engelien ²2002:439).

Generalisierte auditive Agnosien

Die generalisierte auditive Agnosie entsteht durch die bilaterale Läsion der oberen Temporalwindung (*Gyrus temporalis superior*), der Heschl'schen Querwindung (*Gyrus temporalis transversus*) sowie deren afferenten und efferenten Bahnen bis zum Thalamus (*Corpus geniculatum mediale*). Patienten mit generalisierter auditiver Agnosie sind nicht in der Lage, Sinn tragenden Umweltgeräuschen oder gesprochener Sprache eine semantische Bedeutung zuzuordnen (vgl. Engelien 2003:142, ⁵2002:440, Mendez 2001:139, Taniwaki 2000:156, Kazui et al. 1990:476, Lambert et al. 1989:71). Sie können jedoch – von leichten aphasischen Beeinträchtigungen abgesehen – sprechen, schreiben und lesen (vgl. z.B. Mendez 2001:140f).

Reine Geräuschagnosie

Die reine Geräuschagnosie tritt nach bilateralen oder rechtshemisphärischen Läsionen auf, wobei die kritischen Regionen mit denen der generalisierten auditiven Agnosie (siehe oben) übereinstimmen (vgl. u.a. Schnider et al. 1994, Fujii et al. 1990:263, Spreen, Benton & Fincham 1965). Wie das oben beschriebene Patientenklientel sind auch die von Geräuschagnosie Betroffenen bei leichten aphasischen Beeinträchtigungen (mit Ausnahme der Fallbeschreibung von Fujii et al. 1990) in der Lage, gesprochene Sprache zu sprechen, zu schreiben und zu lesen. Zusätzlich verfügen sie über ein in der Regel gut erhaltenes Sprachverständnis. Beeinträchtigungen zeigen sich in der isolierten Fähigkeit, nonverbale Alltags- und Umweltgeräusche zu identifizieren (vgl. Engelien ⁵2002:441, Fujii et al. 1990). Obwohl bisher nur äußerst wenige Fälle beschrieben wurden, ist davon auszugehen, dass die Inzidenz dieser Störung unterschätzt wird. Dies liegt – wie bereits oben erwähnt – daran, dass die Überprüfung der zentral-auditiven Funktionen nicht routinemäßig untersucht wird und dass den Betroffenen ihre Störungen in der Regel unspezifisch erscheinen (vgl. Bauer & Demery ⁴2003:266, Engelien 2003:143), sie also in den seltensten Fällen von Seiten der Patienten verbalisiert werden.

Affektive auditive Agnosie

Die affektive auditive Agnosie gehört wie die im folgenden Unterkapitel beschriebene Phonagnosie zu den *paralinguistischen auditiven Agnosien* und wurde bisher nur nach rechtshemisphärischen Läsionen beschrieben (vgl. Heilman, Scholes & Watson 1975:69). Bei diesen Agnosien ist die Fähigkeit gestört, aus gesprochener Sprache die unterschiedlichen akustischen Parameter mit bestimmten para- oder metalinguistischen Informationen zu verbinden. Wie bei der generalisierten auditiven und der reinen Geräuschagnosie sind Sprechen, Lesen und Schreiben weitestgehend unbeeinträchtigt. Die Inhalte gesprochener Sprache werden erfasst, jedoch sind die Betroffenen nicht in der Lage, beispielsweise das Geschlecht und das ungefähre Alter eines Sprechers oder die emotionale Konnotation von Äußerungen zu verstehen (vgl. Engelen ⁵2002:442).

7.3.2. Phonagnosie

Mehrere Untersuchungen von Van Lancker et al. (Van Lancker, Kreiman & Cummings 1989, Van Lancker et al. 1988, Van Lancker & Kreiman 1987, Van Lancker & Canter 1982) beschreiben eine weitere Form der paralinguistischen Agnosie, die *Phonagnosie*². Die von ihr, sowie inzwischen von zahlreichen weiteren Autoren (siehe unten), beschriebenen Patienten zeigen deutliche Defizite im Bereich des Erkennens bzw. Diskriminierens von Stimmen. Die folgenden Abschnitte stellen überblicksweise eine Auswahl der bisherigen Studien zur Phonagnosie vor.

Wahrnehmungsexperimente zur Phonagnosie

In ihrer Studie von 1982 untersuchten Van Lancker & Canter neun nicht-aphasische Patienten (Alter = 61,7 Jahre, SD = 16,2) mit einer Läsion der rechten (RBD)³ und 21 aphasische Patienten (Alter = 52,4 Jahre, SD = 15,3) mit einer Läsion der linken Hemisphäre (LBD)⁴. Dabei bedienten sie sich einer 4-AFC-Aufgabe, um die Identifizierungsfähigkeit der Patienten für berühmte Gesichter bzw. Stimmen zu überprüfen (vgl. Van Lancker & Canter 1982:189f). Die Auswahl der berühmten Gesichter bzw. Stimmen wurde mit Hilfe zweier unterschiedlicher, zerebral gesunder, Personengruppen (N = 44 bzw. 24) getroffen, die mit dem entsprechenden Stimulus (Fotografie oder Stimmaufnahme) konfrontiert wurden und aus einer Auswahl von fünf Namen

²In Anlehnung an die Prosopagnosie, die sich des griechischen Wortes *prosōpon* (= Gesicht) bedient, schlagen Van Lancker & Canter den Terminus *Phonagnosie* (vom griechischen *phonē* (= Stimme)) vor (vgl. Van Lancker & Canter 1982:188).

³*Right-Brain-Damaged*.

⁴*Left-Brain-Damaged*.

7. Neurolinguistische und -psychologische Grundlagen

den der entsprechenden Persönlichkeit auswählen mussten. Sieben Fotografien wurden von nur maximal einer Person falsch identifiziert und bildeten somit das Stimuluset für die Aufgabe zur Gesichteridentifizierung. Mit Hilfe einer ähnlichen Prozedur ($N = 24$) wurden sieben Stimmproben mit einer Länge von 10 bis 15 s ausgewählt, deren Inhalte keine Rückschlüsse auf den Sprecher zuließen. In der Testsituation sollten die teilnehmenden LBD- und RBD-Patienten dem Stimulus (Fotografie bzw. Stimmaufnahme) jeweils einen von vier geschriebenen und vorgelesenen Namen zuordnen. Die Ergebnisse zeigt Tabelle 7.1.

| korr. Items | Anzahl der Probanden | | | |
|-------------|----------------------|---------|-------------|---------|
| | LBD (N = 21) | | RBD (N = 9) | |
| | Gesichter | Stimmen | Gesichter | Stimmen |
| 7 | 12 | 10 | 3 | 4 |
| 6 | 8 | 8 | 2 | – |
| 5 | – | 2 | – | – |
| 4 | 1 | 1 | 3 | 1 |
| 3 | – | – | 1 | 1 |
| 2 | – | – | – | 2 |
| 1 | – | – | – | – |
| 0 | – | – | – | – |
| | 6,48 | 6,29 | 5,33 | 4,89 |

Tabelle 7.1.: Leistungen im Bereich der Gesichter- bzw. Stimmidentifizierung

Auffällig ist hierbei die relativ stärkere Streuung innerhalb der (kleineren) Gruppe der RBD-Patienten. Während die LBD-Patienten durchschnittlich 6,48 ($SD = ,75$) bzw. 6,29 ($SD = ,85$) der Gesichter- bzw. Stimmidentifizierungsaufgaben korrekt lösten, gelang dies den RBD-Patienten nur bei 5,33 ($SD = 1,58$) bzw. 4,89 ($SD = 2,21$) Aufgaben. Für eine zweite Analyse der Daten definierten Van Lancker & Canter die Erkennungsleistung als pathologisch, wenn weniger als fünf der sieben Aufgaben korrekt gelöst werden konnten. Dies war nur bei einem der 21 aphasischen LBD-Patienten (4,8 %), aber bei vier der neun nicht-aphasischen RBD-Patienten (44,4 %) ($p = .019$) der Fall (vgl. Tabelle 7.2, Van Lancker & Canter 1982:191, siehe auch Assal et al. 1976). Drei der vier RBD-Patienten zeigten darüber hinaus Defizite bei der Gesichteridentifizierung, was den Verdacht nahe legt, dass beiden Auffälligkeiten eine gemeinsame Dysfunktion zugrunde liegen könnte. Weiter zeigen die Ergebnisse,

7. Neurolinguistische und -psychologische Grundlagen

dass das Identifizieren einer Stimme eine Fähigkeit ist, die nicht an die sprachlichen Leistungen der linken Hemisphäre gekoppelt ist und möglicherweise auch nicht mit einer gleichzeitigen reinen Geräuschagnosie einhergehen muss (vgl. Van Lancker & Canter 1982:193).

| | <i>face recognition</i> | | <i>voice recognition</i> | |
|--------------------|-------------------------|----------------|--------------------------|----------------|
| | LBD (N = 21) | RBD (N = 9) | LBD (N = 21) | RBD (N = 9) |
| Pass (5-7 correct) | 20 | 5 | 20 | 5 |
| Fail (0-4 correct) | 1 | 4 | 1 | 4 |
| | $p = .019$ | | $p = .019$ | |

Tabelle 7.2.: Leistungen im Bereich der Gesichter- bzw. Stimmidentifizierung

Die Aussagekraft der Ergebnisse muss möglicherweise eingeschränkt werden, da sich das Durchschnittsalter der beteiligten Probanden deutlich voneinander unterscheidet (61,7 Jahre der LBD-Patienten vs. 52,4 Jahre der RBD-Patienten). Auf den ersten Blick mag möglicherweise auch die unterschiedlich hohe post-onset-time (2,0 vs. 8,9 Monate) einen Einfluss auf die Erkennungsleistungen haben, was von den Autoren der Studie jedoch überprüft und ausgeschlossen werden konnte (vgl. Van Lancker & Canter 1982:192). Darüber hinaus wurden Patienten mit unterschiedlichen Ätiologien in der Studie eingesetzt (nicht näher definierte zerebrovaskuläre Infarkte bzw. Schädel-Hirn-Traumen) und auch das Klassifikationsinstrument für die aphasischen Patienten bleibt unklar.

Die Ergebnisse zweier weiterer Studien von Van Lancker & Kreiman (1987) bzw. von Van Lancker et al. (1988) legen nahe, dass die bereits im Rahmen der Einführung (vgl. Kapitel 2 auf Seite 4) erläuterten Termini *Stimmdiskriminierung* und *Sprecheridentifizierung* zwei verschiedene Leistungen beschreiben, denen unterschiedliche kognitive Prozesse zugrunde liegen. An der Untersuchung nahmen insgesamt 45 Patienten mit einer uni- oder bilateralen Läsion sowie 48 zerebral gesunde Kontrollpersonen teil. Zur Überprüfung der Fähigkeit zur *Stimmidentifizierung* wurden die Patienten und Probanden gebeten, im Rahmen einer 4-AFC-Aufgabe, ein viersekündiges Stimmuster einer Fotografie bzw. dem Namen einer berühmten Persönlichkeit zuzuordnen. Nach der Bearbeitung aller Testitems wurden die Probanden für jeden einzelnen Zielsprecher gefragt, ob sie dessen Stimme unter normalen Umständen (er-)

7. Neurolinguistische und -psychologische Grundlagen

kennen würden. Lediglich die so als „bekannt“ definierten Stimmen wurden bei der Datenauswertung berücksichtigt. Für die Überprüfung der Fähigkeit der *Stimmdiskriminierung* mussten die Teilnehmer 26 Stimmpaare mit „gleich“ oder „ungleich“ bewerten. Bei der Analyse der Messwerte zeigten sich drei Argumente, die für eine Differenzierung der genannten Leistungen, *Identifizierung* vs. *Diskriminierung*, sprachen und hier kurz skizziert werden sollen:

1. Die Leistungen in der Diskriminierungs- und Identifizierungsaufgabe korrelierten bei der Gruppe der Kontrollprobanden nur wenig miteinander ($r = .41$, $p = .01$), bei der LBD- bzw. LBD-Gruppe gar nicht ($r = .20$, *n.s.*) (vgl. Van Lancker & Kreiman 1987:831).
2. Sowohl die LBD-, als auch die RBD-Patienten zeigten Beeinträchtigungen bei der Stimmdiskriminierung, während lediglich die RBD-Patienten auch bei der Identifizierungsaufgabe signifikant schlechter abschnitten (vgl. Tabelle 7.3).
3. 44 % der zerebral geschädigten Patienten (14 von 32) wiesen einen Leistungsunterschied zwischen den beiden Aufgabentypen von mehr als zwei Standardabweichungen auf. Im Gegensatz dazu zeigte nur einer der 48 Kontrollprobanden eine solch große Diskrepanz (vgl. Van Lancker & Kreiman 1987:831).

| Aufgabe | Kontrollgr. | LBD | RBD |
|--------------------------|--------------|--------------|--------------|
| Sprechererkennung | 82,1 | 81,8 | 62,9 |
| Range | 46,7-100 | 68,8-93,3 | 16,0-95,5 |
| <i>S.D.</i> | <i>11,39</i> | <i>7,93</i> | <i>20,00</i> |
| Diskriminierung | 87,2 | 76,4 | 69,9 |
| Range | 53,8-100 | 53,8-96,2 | 42,3-92,3 |
| <i>S.D.</i> | <i>9,96</i> | <i>13,84</i> | <i>16,22</i> |

Anm.: alle Angaben in %.

Tabelle 7.3.: Ergebnisse der Sprechererkennungs- und Diskriminierungsaufgabe

Aufgrund der Testung und des Vergleichs von Patienten mit sehr heterogenen Ätiologien (Schlaganfall mit nicht näher angegebener Genese, Z.n. Kraniotomie sowie Raumfordernde Veränderungen wie Blutungen und Meningeome) ist die Vergleichbarkeit und Aussagekraft der Daten möglicherweise eingeschränkt. Darüber hinaus

7. Neurolinguistische und -psychologische Grundlagen

muss kritisch hinterfragt werden, ob die Selbst-Beurteilung der eigenen Wiedererkennungsfähigkeit einer berühmten Stimme gerade für globalaphasische Patienten sowie für hirngeschädigte Patienten mit unklaren neuropsychologischen Begleiterscheinungen überhaupt zuverlässig möglich ist.

Vergleichbare Ergebnisse wie die soeben geschilderten ergab die Untersuchung zur Stimmverarbeitung von Van Lancker et al. (1988), die die Autoren mit Hilfe des bereits oben beschriebenen Testdesigns mit sechs BD-Patienten durchführten (zwei LBD-Patienten, ein RBD-Patient und drei Patienten mit bilateralen Läsionen (BBD)⁵). Tabelle 7.4 gibt einen Überblick über die Erkennungsleistungen der Patienten und die Lokalisation der Hirnschädigung.

| Patient | Läsion | Leistung (%) | |
|---------|-----------------------------|-----------------|-----------------|
| | | Diskriminierung | Wiedererkennung |
| RBD | tempoparietal | 58 | 41 |
| LBD | fronto-tempo-parietal | 58 | 84 |
| LBD | temporal | 69 | 92 |
| BBD | R temporal, L tempoparietal | 50 | 92 |
| BBD | R parietal, L parietal | 73 | 16 |
| BBD | R parietal, L okzipital | 58 | 41 |
| KG | - | 85 | 83 |

Tabelle 7.4.: Ergebnisse der Stimmverarbeitungsaufgaben

Diese ersten Hinweise auf neuroanatomische Korrelate der Phonagnosie nahmen Van Lancker, Kreiman & Cummings (1989) zum Anlass einer weiteren Untersuchung. Wie bereits in den vorherigen Studien nutzten Van Lancker und ihre Kollegen erneut das bereits erläuterte Testdesign. Dazu führten sie die Stimmverarbeitungsaufgaben (Identifizieren und Diskriminieren) mit 44 BD-Patienten (23 aphasische LBD-, 15 RBD-, 6 BBD-Patienten)⁶ und zusätzlich eine isolierte Identifizierungsaufgabe mit 12 weiteren BD-Patienten durch (2 aphasische LBD-, 10 RBD-Patienten), so dass insgesamt 56 zerebral geschädigte Patienten untersucht wurden. Für die Diskriminierungsaufgabe wurden 13 gleiche bzw. 13 ungleiche Stimmpaare eingesetzt, die aus den Äußerungen von zehn unterschiedlichen Sprechern extrahiert wurden.

⁵ *Bilateral-Brain-Damaged*.

⁶ Inklusive der 15 LBD-, 11 RBD- und 6 BBD-Patienten aus Kreiman & Van Lancker 1988.

7. Neurolinguistische und -psychologische Grundlagen

Die 25 Stimmproben für die Identifizierungsaufgabe stammten von berühmten Persönlichkeiten. Dazu wurden den Teilnehmern zu jedem präsentierten Stimmuster vier Fotografien (Zielperson + drei Distraktoren, 4-AFC, *closed set*) mit Namen zur Auswahl gestellt. Die Kontrollprobanden mussten den korrekten Namen aus einer Auswahl von fünf Möglichkeiten (5-AFC, *closed set*) treffen. Auch bei dieser Studie wurden die Teilnehmer nach der Bearbeitung sämtlicher Items zu jeder Persönlichkeit befragt, ob sie deren Stimme erkennen würden. Nur die als „bekannt“ bezeichneten Stimmen wurden bei der Auswertung der Daten berücksichtigt.

Die Ergebnisse der Aufgaben sind in Tabelle 7.5 aufgeführt (vgl. Van Lancker, Kreiman & Cummings 1989:668).

Eine Analyse der Messwerte zeigt, dass sich die Leistungen der LBD-Patienten und der Kontrollpersonen im Bereich der Stimmentifizierung statistisch nicht voneinander unterscheiden. Hochsignifikante Unterschiede zeigen sich jedoch zwischen den Leistungen der RBD-Gruppe und denen der Kontroll- ($p < ,001$) bzw. der LBD-Gruppe ($p < ,001$). Bei der Diskriminierungsaufgabe unterscheiden sich sowohl die Leistungen der LBD- ($p < ,01$), als auch die der RBD-Gruppe ($p < ,01$) signifikant von denen der Kontrollgruppe (vgl. Van Lancker, Kreiman & Cummings 1989:668). Damit bestätigen sie die Ergebnisse von Van Lancker & Kreiman 1988. Auch hier tritt eine beeinträchtigte Stimmentifizierung mit einer Läsion der rechten Hemisphäre zusammen, wohingegen eine Beeinträchtigung der Stimmdiskriminierungsfähigkeit mit einer Störung der rechten *oder* linken Hemisphäre auftritt.

Die von insgesamt 43 Patienten vorliegenden CT-Scans wurden daraufhin hinsichtlich der intrahemisphärischen Läsionen und möglicher Korrelationen mit den Leistungen in den spezifischen Aufgaben analysiert. Dabei zeigte sich eine signifikante Korrelation zwischen einer Schädigung des rechten Parietallappens und einem Defizit bei der Stimmwiedererkennung. Eine seitenunabhängige Schädigung des Temporallappens ging hingegen mit einer beeinträchtigten Diskriminierungsleistung einher (vgl. Van Lancker, Kreiman & Cummings 1989:665, 668). Dass Identifizieren und Diskriminieren zwei unabhängige Leistungen sind, zeigt die doppelte Dissoziation, die die

| Aufgabe | Gruppe | | |
|---------------------|--------|---------|----------------|
| | LBD | RBD | Kontrollgruppe |
| Identifizierung (%) | 80,10 | 58,57** | 82,10 |
| Diskriminierung (%) | 77,07* | 68,21** | 87,19 |

Tabelle 7.5.: Ergebnisse der Stimmverarbeitungsaufgaben 2

7. Neurolinguistische und -psychologische Grundlagen

Ergebnisse der Untersuchung dokumentieren: sieben der 43 BD-Patienten erzielten zwar gute Ergebnisse bei der Diskriminierungsaufgabe, lagen jedoch bei der Identifizierung mehr als zwei Standardabweichungen unter den Durchschnittsleistungen der Kontrollgruppe. Eine gleich große Diskrepanz (> 2 SD) zeigten neun Patienten beim Diskriminieren von Stimmen, die jedoch bei der Identifizierungsaufgabe sehr gute Ergebnisse erzielten (vgl. Van Lancker, Kreiman & Cummings 1989:672).

Unklar bleibt, weshalb die Autorinnen die Leistungen der BD-Gruppen im Rahmen einer 4-AFC-Aufgabe, die Leistungen der Kontrollgruppe hingegen mit Hilfe einer 5-AFC-Aufgabe erheben. Eine Vergleichbarkeit der Ergebnisse ist daher fraglich und anzuzweifeln. Ebenso muss – wie bereits weiter oben – hinterfragt werden, ob die Urteile über die individuelle Bekanntheit einer berühmten Stimmen von globalaphasischen und / oder neuropsychologisch beeinträchtigten Patienten als zuverlässige Informationsquelle angenommen werden können.

Inwiefern nun die Ergebnisse bildgebender Verfahren mit denen der oben beschriebenen Wahrnehmungsexperimente übereinstimmen, ist Inhalt des folgenden Abschnittes.

Ergebnisse bildgebender Untersuchungsverfahren

Um die neuronalen Substrate, die an der Verarbeitung stimmlicher Signale beteiligt sind, identifizieren zu können, nutzen einige Studien die unterschiedlichen Methoden der funktionellen Bildgebung, wie z.B. die Positronen-Emissions-Tomographie (PET) oder die funktionelle Magnetresonanztomographie (fMRT). Beide gehören zu den so genannten *metabolischen Verfahren*, die die neuronale Aktivität indirekt über Veränderungen des Stoffwechsels messen. Diese finden sich in unmittelbarer Nachbarschaft aktivierter Neuronenpopulationen und bewirken eine regionale zerebrale Blutflusssteigerung (*regional cerebral blood flow, rCBF*) (vgl. Hartje & Poeck ⁵2002:52).

Die ersten, die die PET nutzten, um die durch Stimmverarbeitung indizierten zerebralen Aktivierungsmuster zu untersuchen, waren Imaizumi und Kollegen (vgl. Imaizumi et al. 1997). Dazu wurden sechs gesunde Probanden mit vier unterschiedlichen, initial unbekannt, Stimmen familiarisiert (S1 - S4). Jeder Proband sollte u.a. jedem der 192 Stimuli (Einzelwörter) eine der zuvor trainierten Stimmen zuordnen. Dies erfolgte, wie bei fMRT-Untersuchungen üblich, mit Hilfe eines Tasters. Dazu wurden vier Fingern je ein Sprecher zugeordnet (Zeigefinger - Sprecher 4, Mittelfinger - Sprecher 3 usw.). Eine Analyse der Scans zeigte insbesondere eine – mit heutigem Wissen erwartungsgemäße – Aktivierung der temporalen Pole beider Hemisphären. Die Autoren bringen dies mit der Rolle dieser Region für das Wiedererkennen von Objekten

7. Neurolinguistische und -psychologische Grundlagen

und entsprechenden Gedächtnisleistungen in Verbindung (vgl. Imaizumi 1997:2811). Darüber hinaus berichten sie von einer signifikanter Aktivierung des rechten *Gyrus lingualis*, der an der Geschlechterzuweisung beteiligt ist, die für das Erkennen eines Sprechers relevant sein kann (vgl. Imaizumi 1997:2012, siehe dazu auch Lattner, Meyer & Friederici 2005). Die Zuordnung der Finger zu bestimmten Sprechern erscheint – obwohl es sich um zerebral gesunde, junge Probanden handelt – umständlich und kompliziert und lenkt, da die Reaktion „as soon as possible“ erfolgen soll, von der eigentlichen Stimmerkennungsaufgabe ab.

Nakamura et al. (2001) untersuchten im Rahmen ihrer PET-Studie die Veränderungen des rCBF bei neun gesunden männlichen Probanden im Rahmen dreier Aufgaben, die sich in der Art der Stimuli voneinander unterschieden. Die Stimuli der „*familiar*-Aufgabe“ bestanden zu drei Viertel aus den Stimmen unbekannter Sprecher ($N > 50$) und zu einem Viertel aus den Stimmen von Freunden und Kollegen. Die Stimuli der „*Vokal*-Aufgabe“ bestanden komplett aus fremden Stimmen. Bei einem Viertel der Sätze bestand das erste Phonem aus einem Vokal, bei drei Viertel aus einem Konsonant. Die Stimuli der „*self*-Aufgabe“ bestanden zu einem Viertel aus der eigenen Stimme und zu drei Viertel aus fremden Stimmen (siehe oben). Bei jedem Testitem wurden die gesprochenen Sätze (2 s) zweimal präsentiert und der Proband sollte per Knopfdruck entscheiden, ob die Stimme bekannt oder unbekannt (*familiar*-Aufgabe) war, der Satz mit einem Vokal begann oder nicht (*Vokal*-Aufgabe) bzw. der Satz von der eigenen oder einer fremden Stimme gesprochen worden war (*self*-Aufgabe). Eine Auswertung des rCBF zeigte, dass die geforderte Leistung der Stimmidentifizierung den frontalen Pol, den rechten temporalen Pol, den rechten entorhinalen Kortex und den linken Prucuneus aktivierte. Dies legt nahe, dass temporale, parietale und frontale Regionen ein Netzwerk bilden, das an dem Erkennen von Stimmen beteiligt ist (vgl. Nakamura et al. 2001:1050). Dies wiederum unterstützt die Ergebnisse der neuropsychologischen Wahrnehmungsexperimente, die von einer eingeschränkten Stimmerkennung bei Patienten mit temporalen Läsionen berichteten (vgl. Scott et al. 1997, Kapur et al. 1992, 1994, Ellis, Young & Critchley 1989). Auffallend ist hierbei, dass alle der genannten Regionen an den Funktionen des episodischen Gedächtnisses beteiligt sind.

Während Van Lancker et al. (1988) und Van Lancker, Kreiman & Cummings (1989) eine Beteiligung der rechten parietalen Region beim Identifizieren und die bilateralen temporalen Regionen beim Diskriminieren von Stimmen postulieren, zeigen die Probanden von Nakamura et al. (2001) in der „*familiar*-Aufgabe“ Aktivitäten in einigen eng umschriebenen Regionen des rechten Temporallappens (temporaler Pol und entorhinaler Kortex, siehe oben). Ein Vergleich der Ergebnisse erscheint auf Grund der

7. Neurolinguistische und -psychologische Grundlagen

unterschiedlichen Testdesigns jedoch unzulässig.

Die PET-Scans während der „*self*-Aufgabe“ zeigten eine bilaterale Aktivierung im frontalen Kortex, die sich von denen der „*familiar*-Aufgabe“ unterschieden: der Anstieg des rCBF im rechten inferioren frontalen Sulcus und im parainsularen Kortex war nun größer als in der „*familiar*-Aufgabe“, so dass die Autoren daraus schließen, dass diese Regionen an dem Erkennen der eigenen Stimme beteiligt sein könnten (vgl. Nakamura et al. 2001:1051). Bemerkenswert ist, dass der signifikante Anstieg des rCBF während der „*familiar*-Aufgabe“ im rechten temporalen und im linken frontalen Pol mit der Anzahl der korrekt identifizierten Stimmen korrelierte. Nakamura et al. (2001) folgern daraus, dass diese beiden Regionen unmittelbar daran beteiligt sind, eine gehörte Stimme mit bekannten Stimmen im Personengedächtnis zu vergleichen.

Dass der temporale Pol an den Leistungen des Personen- / episodischen Gedächtnisses und an weiteren Erkennungsleistungen (z.B. Gesichter, Gebäude) beteiligt sind, belegen zahlreiche weitere Studien (vgl. u.a. Ellis, Young & Critchley 1989, Kapur et al. 1992, Markowitsch 2003, 1995, Markowitsch et al. 2000, Nakamura & Kubota 1996). Der linke frontale Pol, der während der „*familiar*-Aufgabe“ die größte Aktivität zeigte, spielt ebenfalls eine Rolle beim Finden von Inhalten im episodischen Gedächtnis (vgl. z.B. den Überblick von Cabeza & Nyberg 2000:31ff).

Belin & Zatorre (2003) untersuchten mit Hilfe der fMRT 14 gesunde Erwachsene (9 w, 5 m) während diese mit insgesamt 24 Stimuliblöcken (bestehend aus je 12 Einsilbern, Dauer: 20 s) beschallt wurden. In Anlehnung an das Adaptions-Paradigma⁷ wurde eines von zwei akustischen Merkmalen symmetrisch manipuliert, während das andere konstant blieb. Bei der Bedingung „adapt - speaker“ bestanden die Blöcke aus 12 unterschiedlichen Einsilbern (z.B. had, hod, hawed, head, heard etc.) die von ein und demselben Sprecher vorgetragen wurden. Die linguistische Information variierte also maximal, wohingegen die sprecherspezifische Information wiederholt wurde. Bei der zweiten Bedingung („adapt - syllable“) wurde ein und derselbe Einsilber von 12 verschiedenen Sprechern gesprochen. Bei dieser Bedingung variierten entsprechend die sprecherspezifischen Charakteristika maximal, während der linguistische Inhalt des Gesprochenen konstant blieb (vgl. Belin & Zatorre 2003:2105f). Verglichen mit der Ruhebedingung zeigten die Stimuliblöcke die zu erwartende starke bilaterale Aktivierung großer Teile der superioren temporalen Gyri um die Heschl'sche Querwin-

⁷Das Paradigma basiert auf der Eigenschaft von Neuronenpopulationen, ihre Feuerungsrate als Reaktion auf einen wiederholten Stimulus zu reduzieren.

7. Neurolinguistische und -psychologische Grundlagen

dung. Der direkte Vergleich der beiden beschriebenen Bedingungen („adapt - speaker“ vs. „adapt - syllable“) zeigte einen signifikanten Unterschied des BOLD-Signals⁸ im anterioren Anteil des superioren temporalen Gyrus der rechten Hemisphäre, also in der Nähe des rechtsseitigen *Sulcus temporalis superior (STS)*. Der STS zeigte eine signifikant ($p < ,05$) geringere Aktivität bei der „adapt - speaker“-Bedingung als bei dem „adapt - syllable“-Paradigma. Die Aktivität nahm also ab, wenn eine Einzelstimme gehört wurde (vgl. Belin & Zatorre 2003:2107). Dies kann in Anlehnung an das bereits oben genannte Adaptions-Paradigma damit erklärt werden, dass die kritischen Neuronenpopulationen aufgrund des Einzelsprechers und somit der Wiederholung der charakteristischen, idiosynkratischen Sprechermerkmale ihre Feuerungsrate reduzieren.

Mehrere Autoren teilen die Meinung, dass das Wiedererkennen von Stimmen in Analogie zur Wiedererkennung von Gesichtern oder Objekten organisiert sein könnte (vgl. Belin, Fecteau & Bédard 2004:129, Review von Mesulam 1998), eine Stimme also als „auditory face“ (Belin, Fecteau & Bédard 2004:129) interpretiert werden kann. Ein weit verbreitetes Modell zur Erkennung von Gesichtern, Objekten und Wörtern, das im Abschnitt *Modelle der Stimmverarbeitung* ab Seite 99 vorgestellt wird, entwickelten u.a. Bruce & Young (vgl. 1988:318). Dieses Modell sieht so genannte *face recognition units (FRU)* vor, in denen jene Merkmale eines Gesichtes gespeichert sind, die als stabil und invariant gelten und auch von wechselnden Sichtwinkeln und Lichtverhältnissen unbeeinflusst bleiben (vgl. Burton, Bruce & Johnston 1990:361). Nimmt man analog dazu die Existenz entsprechender *voice recognition units (VRU)* an, so könnten diese nach Ansicht von Belin & Zatorre im rechten anterioren STS lokalisiert sein (vgl. Belin & Zatorre 2003:2108). In einer früheren Studie konnten Belin, Zatorre & Ahad (2002) außerdem zeigen, dass einige Anteile des rechten STS die einzigen Regionen des auditorischen Kortexes waren, die diese Aktivierung auch bei nicht-linguistischen Vokalisationen wie Lachen, Schreien und Räuspern zeigten (vgl. Belin, Zatorre & Ahad 2002:17).

Die bisherigen Ergebnisse legen nahe, dass die anterioren Anteile des superioren Temporallappens, insbesondere die der rechten Hemisphäre, an der Extraktion paralinguistischer Merkmale und somit auch am Wiedererkennen von Stimmen beteiligt sind (vgl. Belin & Zatorre 2003:2108, Crinion et al. 2003, Belin, Zatorre & Ahad 2002).

Dies bestätigen auch die Ergebnisse einer fMRT-Studie von Von Kriegstein & Gi-

⁸ *Blood Oxygenation Level Dependent Contrast*

7. Neurolinguistische und -psychologische Grundlagen

raud (2004). Sie postulieren, dass einige isolierte kortikale Regionen entlang des STS an unterschiedlichen Aspekten der Stimmverarbeitung beteiligt sind und dass diese im Rahmen der Sprechererkennung miteinander Netzwerke bilden. Zuvor konnten sie in einer früheren Studie (Von Kriegstein et al. 2003) eine größere Aktivierung des rechten anterioren STS und eines Teils des rechten Precuneus nachweisen, wenn Probanden einen Sprecher unabhängig vom semantischen Inhalt des Stimulussatzes erkennen sollten. Bestand die Aufgabe jedoch darin, einen bestimmten Inhalt unabhängig vom Sprecher zu erkennen, so zeigte der mittlere Anteil des STS eine größere Aktivierung (vgl. Von Kriegstein et al. 2003:48). Von Kriegstein & Giraud konfrontierten dazu ihre Probanden (4 w, 5 m) mit einem identischen Set von Sätzen, während diese entweder einen Zielsatz unabhängig des Sprechers oder einen Sprecher unabhängig des gesprochenen Satzes erinnern sollten. Die eine Hälfte der Stimmen stammte dabei von bekannten Sprechern, die andere Hälfte von unbekanntem (jeweils $N = 14$). Die Messungen zeigten eine größere Aktivierung des rechten anterioren und posterioren STS bei der Sprechererkennungsaufgabe als bei dem Erkennen des semantischen Inhalts. Während der rechte anteriore STS bei beiden Stimmkategorien Aktivität zeigte, war dies bei dem rechten posterioren STS nur bei den unbekanntem Sprechern der Fall (vgl. Von Kriegstein & Giraud 2004:951ff). Eine Analyse der Netzwerkverbindungen zeigte, dass während der Sprechererkennung der anteriore und posteriore Teil des STS mit einer Region interagiert, der die Verarbeitung akustischer Stimmmerkmale zugeschrieben wird und sich im mittleren / anterioren Teil des STS befindet. Die Autorinnen schließen daraus, dass drei Regionen des STS für die unterschiedlichen Aspekte der Stimmverarbeitung zuständig sind und deren Interaktion untereinander mit dem Grad der Stimmvertrautheit variiert (vgl. Von Kriegstein & Giraud 2004:954).

Lattner, Meyer & Friederici (2005) untersuchten mit Hilfe der fMRT u.a. die neurophysiologische Verarbeitung der für eine erfolgreiche Sprechererkennung relevanten akustischen Parameter. Dazu wurden 16 Probanden (8 w, 8 m) gebeten, die Natürlichkeit einer Stimme zu beurteilen. Die Ergebnisse der Scans legen nahe, dass unterschiedliche Teile des rechten superioren temporalen Kortexes auch unterschiedliche akustische Parameter verarbeiten. Die Autoren konnten zeigen, dass die Tonhöhe des Sprachsignals in den Regionen um die Heschl'sche Querwindung verarbeitet werden, die spektrale Signalinformation in den posterioren Anteilen des superioren temporalen Gyrus (STG) und in den rechts- und linkshemisphärischen Arealen des *Planum parietale*. Die charakteristischen, prototypischen Informationen einer Stimme werden offenbar in den anterioren Anteilen des rechten STG verarbeitet (vgl. Lattner, Meyer & Friederici 2005:15f).

7. Neurolinguistische und -psychologische Grundlagen

Neben den genannten Untersuchungen existieren noch zahlreiche weitere, die u.a. der Frage der Personenerkennung, also der Integration von visuellen (Gesicht) und auditiven (Stimme) Informationen, nachgehen. Shah et al. (2001) maßen mit Hilfe der fMRT die neuronalen Antworten auf diese beiden Modalitäten. Dazu wurden Probanden parallel mit den Gesichtern und Stimmen bekannter bzw. unbekannter Personen konfrontiert. Unabhängig von der Bekanntheit des Stimulus konnten bei den Gesichtern bilaterale Aktivierungen im *Gyrus fusiformis* („fusiform face area“, Kanwisher, McDermott & Chun 1997:4302) bzw. bei Stimmen – ebenfalls bilateral – in den superioren temporalen Gyri demonstriert werden (siehe dazu auch die Studie von Steeves et al. 2006). Sowohl bei den bekannten Gesichtern als auch den bekannten Stimmen zeigten sich modalitätsübergreifend neuronale Aktivitäten im *Splenium* und dem posterioren *Gyrus cinguli*, der zusammen mit dem *Hippocampus* den wichtigsten Teil des limbischen Systems bildet. Die Autoren schließen daraus, dass beim Wiedererkennen einer Person Informationen von modalitätsspezifischen Modulen im temporalen Kortex zum retrosplenialen Kortex fließen. Diese Region ist maßgeblich am episodischen Gedächtnis beteiligt und könnte eine Schlüsselfunktion bei der Einschätzung der Bekanntheit einer Person spielen (vgl. Shah et al. 2001:804).

Zusammengefasst legen die Untersuchungsergebnisse nahe, dass unterschiedliche kortikale Areale an der Verarbeitung unterschiedlicher Stimminformationen beteiligt sind, wobei den Regionen entlang des *Sulcus temporalis superior* eine besondere Bedeutung zukommt. Darüber hinaus unterstützen sie ein funktionales Modell, in dem linguistische, affektive und identitätsrelevante Informationen auf unterschiedlichen Bahnen verarbeitet werden und das mit den Modellvorstellungen zur Verarbeitung von Gesichtern in weiten Teilen übereinstimmt.

Modelle der Stimmverarbeitung

„Speaker recognition is probably our most complex auditory cognitive ability, apart from speech perception. Our capacity to rapidly and effortlessly recognize an individual based on novel vocalizations suggests that our auditory system extracts the acoustic features of the vocal signal that present higher inter-individual but little intra-individual variation, to combine them in long-term representations of vocal identities“ (Belin & Zatorre 2003:2105).

7. Neurolinguistische und -psychologische Grundlagen

Um eine theoretische Modellvorstellung zur Verarbeitung stimmlicher Signale zu entwickeln, müssen die gewonnenen Ergebnisse der verschiedenen Untersuchungen (klinische, fMRT- und PET-Studien) zusammengefasst, ausgewertet und interpretiert werden. Während fMRT- und PET-Studien eher klären, welche Regionen an einer Verarbeitung beteiligt sind, helfen klinische Studien dabei, in Wahrnehmungsexperimenten messbare (Fehl-)Leistungen zu dokumentieren und daraus Schlüsse auf eine mögliche Verarbeitungsstrategie zu ziehen.

So wissen wir aus den Studien von Van Lancker & Kreiman (1987:833) und Van Lancker et al. (1988:205), dass die Fähigkeit Stimmen zu diskriminieren nicht zwangsläufig eine Vorstufe der Stimmerkennung sein muss.

Assal, Aubert & Buttet (1981) konnten zeigen, dass die Verarbeitung von Sprache und die Verarbeitung von Stimmen zwei voneinander unabhängige Leistungen sind. Im Sinne einer doppelten Dissoziation beschrieben die Autoren mehrere Fälle mit erhaltener Sprachverarbeitung und beeinträchtigten Sprechererkennung sowie die Fälle von Aphasikern mit normalen Leistungen bei der Sprechererkennung (Assal, Aubert & Buttet 1981:267f).

Erste Überlegungen, welche kognitiven Prozesse bei der Verarbeitung stimmlicher Reize ablaufen, schildern Fourcin (1968) und Clarke & Becker (1969). Fourcin betont, dass ein Zuhörer dem Signal aufgrund der z.T. erheblichen Intraindividualität von Stimmen zunächst stabile Merkmale entnehmen müsse. Clarke & Becker differenzieren den Wahrnehmungsvorgang in zwei unterschiedliche Teilleistungen. Zum einen in die Leistung, einer Stimme sprecherspezifische Signale zu entnehmen und zum anderen in den Prozess, der eine Zuordnung der extrahierten Merkmale zu einem spezifischen Sprecher ermöglicht.

Entsprechend formuliert auch Brown (1981):

„[...] auditory speaker recognition [is] definable as the human ability to decide on a speaker's identity from hearing a sample of his speech. The paradigm of speaker-recognition experimentation is that the process involves a pattern-matching technique; that is, a representation of the sample uttered by the voice to be recognized (the stimulus voice pattern) is compared with some form of internalized representation of voices previously heard by the listener (reference voice patterns), and a decision is reached on the basis of the degree of similarity between the two.“ (Brown 1981:295)

Ein Modell zur Gesichter-, Objekt- und Worterkennung entwickelten Bruce und Kollegen (vgl. Bruce & Valentine 1985, Valentine & Bruce 1986, Bruce & Young

7. Neurolinguistische und -psychologische Grundlagen

1988:312, 318, 1990:362, Burton, Bruce & Johnston 1990). Dieses Modell zeigt einige Ähnlichkeiten mit Verarbeitungsmodellen anderer Reizklassen, wie z.B. Wörtern (vgl. Morton 1969, Warren & Morton 1982) oder Alltagsobjekte (vgl. Seymour 1979).

Neuner & Schweinberger (2000) kombinierten und erweiterten die bestehenden Modelle von Bruce & Young (1986) und Burton, Bruce & Johnston (1990) zu einem Arbeitsmodell der Personenerkennung (vgl. Abbildung 7.2).

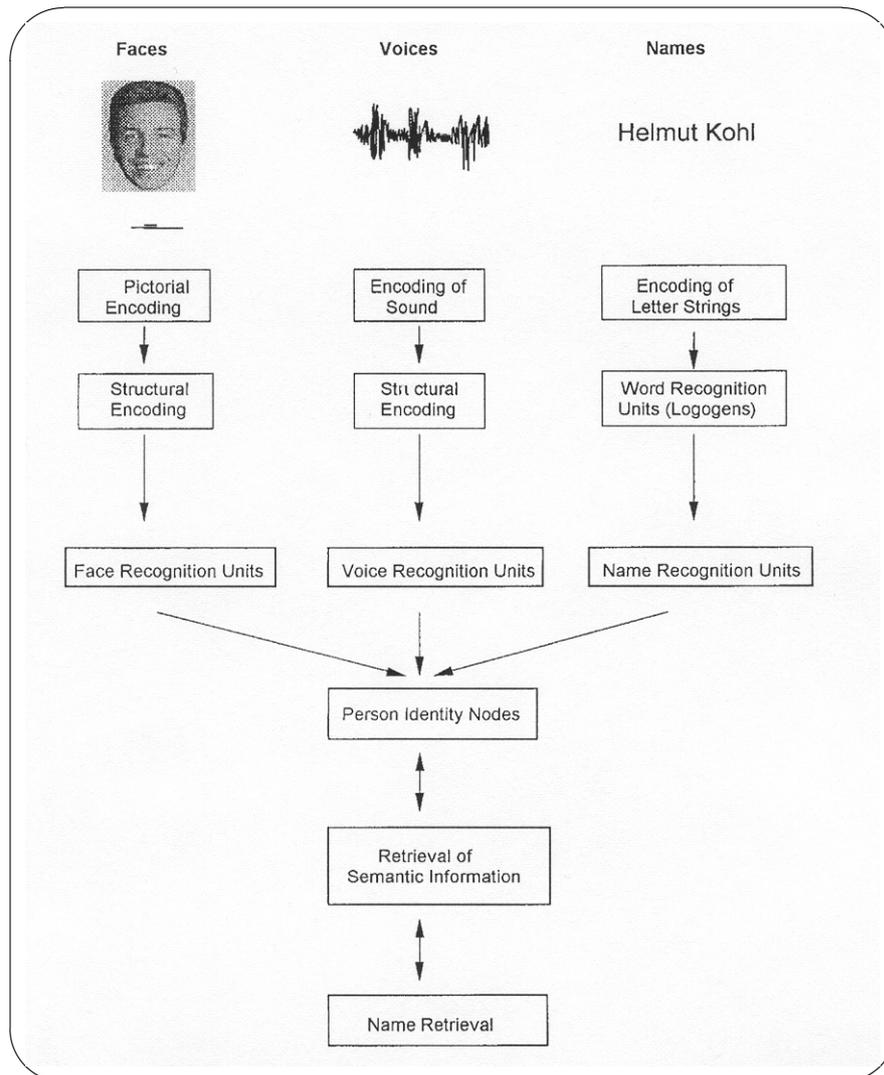


Abbildung 7.2.: Arbeitsmodell zur Personenerkennung
(aus: Neuner & Schweinberger 2000:346)

7. Neurolinguistische und -psychologische Grundlagen

Gemäß dieses Modells wird jeder eingehende Stimulus zunächst einer ersten perceptuellen Analyse unterzogen. Auf der nächsten Verarbeitungsstufe (*Structural Encoding*) werden aus den Stimmen (bzw. Gesichtern) jene Informationen extrahiert, die als stabil und invariant gelten und nicht von situativen Faktoren beeinflusst werden (bei Gesichtern z.B. unabhängig der räumlichen Orientierung oder der Beleuchtung, bei Stimmen unabhängig der Intonation oder des semantischen Inhalts). Diese individuell spezifischen, prototypischen Repräsentationen werden nun mit jenen verglichen, die als *Voice Recognition Units (VRU)*⁹ gespeichert sind. Das Modell geht also davon aus, dass für jeden bekannten Sprecher ein *unique pattern* existiert, das bei Bedarf mit den Merkmalen einer gehörten Stimme verglichen und „abgepasst“ wird. War der Vergleich erfolgreich, so werden die *Person Identity Nodes (PIN)* aktiviert. Die *PIN* bilden eine postperzeptuelle Stufe, auf der die unterschiedlichen Stimulitypen zusammenlaufen und auf der Urteile über die vermeintliche Bekanntheit der Person gefällt werden (vgl. Burton, Bruce & Johnston 1990:361). Die *PIN* sind darüber hinaus die Schnittstellen zum Abrufen weiterer (semantischer) Informationen bzw. des Namens der Person (vgl. Neuner & Schweinberger 2000:345f).

Ein sehr ähnliches Modell schlagen Belin, Fecteau & Bédard (2004) vor (vgl. Abbildung 7.3), dessen einzelne Verarbeitungsstufen sie mit spezifischen zerebralen und zentralnervösen Strukturen in Verbindung setzen.

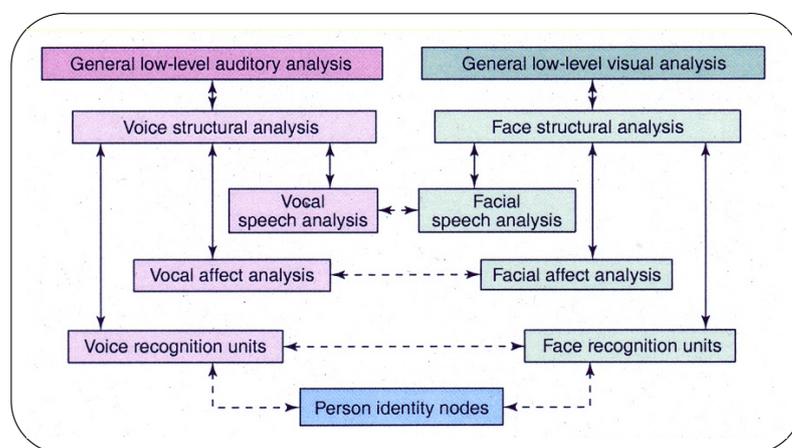


Abbildung 7.3.: Arbeitsmodell zur Stimm- und Gesichtererkennung
(aus: Belin, Fecteau & Bédard 2004:131)

⁹In Anlehnung an die *Face Recognition Units (FRU)* bzw. *Name Recognition Units (NRU)* von Bruce & Young 1986:311, Ellis, Jones & Mosdell 1997, Valentine et al. 1991.

7. Neurolinguistische und -psychologische Grundlagen

Zunächst werden die eingehenden Signale in den subkortikalen Nuclei und den Regionen des primären auditorischen Kortexes einer *Low-Level Analysis* unterzogen. Zuvor durchlaufen die stimmlichen Signale die Stufe des *Structural Encoding*, was möglicherweise bilateral von den Anteilen des mittleren STS, in der Nähe der Regionen des primären auditorischen Kortexes, geleistet werden könnte. Die Verarbeitung der stimmlichen Informationen wird im Anschluss, nach Meinung der Autoren, von drei funktionell unabhängigen Systemen weiter verarbeitet: (1) die Verarbeitung der sprachlichen Informationen im anterioren und posterioren STS und in den inferioren präfrontalen Regionen (überwiegend) der linken Hemisphäre; (2) die stimmlich-affektiven Informationen in den temporo-medialen Regionen, der anterioren Insula, der Amygdala und den inferioren präfrontalen Regionen (überwiegend) der rechten Hemisphäre; (3) die Analyse der stimmlichen Identität unter Einbeziehung der *Voice Recognition Units*, die möglicherweise in den rechten anterioren Regionen des STS verortet sein könnten. Eine abschließende supra-modale Stufe bilden, wie in dem oben erläuterten Modell von Neuner & Schweinberger (2000), die *Person Identity Nodes*.

Anders als bei visuellen Wiedererkennungsaufgaben, bei denen die optischen Stimuli parallel präsentiert werden können (vgl. u.a. Heckmann, Lang & Neundörfer 2001, Heckmann et al. 1994), müssen sie bei auditiven Erkennungsaufgaben nacheinander dargeboten werden. Aus diesem Grunde kann man auch bei dem in der vorliegenden Arbeit verwendeten Testdesign davon ausgehen, dass der Teilnehmer nach jedem gehörten Stimulus einen Abgleich mit dem vertrauten Stimmuster des Ehe- / Lebenspartners vornimmt und in dessen Folge eine bekannt / unbekannt-Entscheidung trifft. Wie bereits weiter oben erläutert, wird eine solche Entscheidung auf der Stufe der *Person Identity Nodes* verortet.

Für die vorliegende Untersuchung soll, in Anlehnung an das von Bruce und Kollegen entwickelte und von Neuner & Schweinberger (2000) erweiterte Modell der Personenerkennung, ein Arbeitsmodell zugrunde gelegt werden, das die im Folgenden erläuterten Stimmverarbeitungsstufen beinhaltet:

1. durch das Nennen des Namens der zu erkennenden Person (*closed set*) wird zunächst das entsprechende *unique pattern* der *VRU* voraktiviert,
2. auf der Verarbeitungsstufe *Encoding of Sound* erfolgt eine Differenzierung in verbale und nonverbale Signalqualität (Sprache vs. Geräusch, Ton, Klang),

7. Neurolinguistische und -psychologische Grundlagen

3. auf der Stufe des *Structural Encoding* separiert der Hörer die intraindividuell stabilen Informationen des Sprechers,
4. vergleicht diese mit dem / den bereits voraktivierten *unique pattern* der VRU und
5. fällt eine bekannt / unbekannt-Entscheidung (*Person Identity Nodes*).

Für eine beeinträchtigte stimmliche Wiedererkennungseistung kämen, nach dem vorgestellten Arbeitsmodell, Störungen auf unterschiedlichen Verarbeitungsstufen infrage. Denkbar wäre z.B. eine Störung auf der Stufe des *Structural Encodings*, die dem Patienten eine Entnahme der relevanten, stabilen Parameter aus dem akustischen Signal nicht mehr oder nur noch eingeschränkt ermöglicht. Auch ein gestörter Zugriff auf die VRU, die die stimmlichen Informationen im Sinne von *unique pattern* beinhalten, oder eine Störung auf Ebene der *Person Identity Nodes* wären für eine beeinträchtigte bekannt / unbekannt-Entscheidung vorstellbar.

Um eine Aussage über die Stimmerkennungseistung zerebral beeinträchtigter Patienten formulieren zu können, sollen in der vorliegenden Untersuchung auch jene Faktoren berücksichtigt werden, die bei den oben vorgestellten Wahrnehmungsexperimenten als kritisch eingestuft wurden:

- es werden ausschließlich Patienten mit *einheitlicher* Ätiologie berücksichtigt,
- die Bekanntheit einer Stimme muss nicht überprüft werden, sondern kann vorausgesetzt werden, da die teilnehmenden Patienten / Probanden sehr gut mit der Zielstimme vertraut sind (Ehe- / Lebenspartner),
- Patienten und Probanden erhalten *identische* Testitems, um eine Vergleichbarkeit zu gewährleisten,
- die Alterszusammensetzung der Patienten- und Probandengruppen sind vergleichbar,
- es wird ein standardisiertes Verfahren zur Diagnostik und Klassifikation der aphasischen Patienten eingesetzt.

Teil II.

Experimenteller Teil

8. Methodik

8.1. Vorüberlegung zur Methodik

Wie bereits in Kapitel 5.1.1 deutlich wurde, sind viele bisherige Untersuchungen zur auditiven Sprecher- / Stimmerkennung als Evaluierungs- oder Identifizierungstests durchgeführt worden. Bei den **Evaluierungs-** bzw. **Rating-Tests** nutzt man semantische Differenzialskalen, anhand derer die Probanden ihre subjektiven Empfindungen psychometrisch ausdrücken können. Dazu gehört u.a. z.B. die Einschätzung *hell vs. dunkel, laut vs. leise, kräftig vs. schwach* etc. Dieses methodische Vorgehen ist für die vorliegende Untersuchung ungeeignet, da die Patienten und Probanden eine subjektive Bewertung interindividuell unterschiedlich wahrgenommener Stimmsignale und kein Stimmerkennen vornehmen.

Die eigentlichen **Stimmerkennungsverfahren** werden als Benenn- oder Matching-Aufgaben durchgeführt. Dies ist in Abbildung 8.1 graphisch dargestellt (in Anlehnung an Bricker & Pruzansky 1976:308).

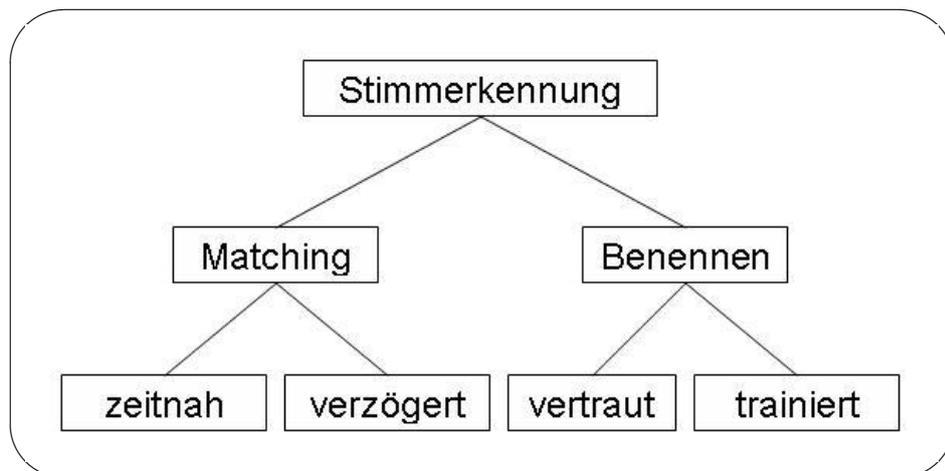


Abbildung 8.1.: Methodenschema der Stimmerkennungsverfahren

Bei einer **Benennaufgabe** sind die Teilnehmer mit der zu identifizierenden Stimme (Zielstimme) durch berufliche und / oder soziale Kontakte vertraut (z.B. Nakamura et al. 2001, Neuner & Schweinberger 2000). Es handelt sich bei den Zielstimmen also um Arbeitskollegen, Familienangehörige oder Freunde. Einige Untersucher setzen auch medial bekannte Stimmen von Politikern, Schauspielern oder Entertainern ein (z.B. Van Lancker et al. 1988, Van Lancker, Kreiman & Wickens 1985). Eine andere Möglichkeit besteht darin, die Probanden mit der Zielstimme vor Beginn der

8. Methodik

Testung zu familiarisieren (z.B. Hollien, Bennett & Gelfer 1983). Das bedeutet, dass die Zielstimme unter kontrollierten oder unkontrollierten Bedingungen aus Anlass der Untersuchung trainiert wird. Das Ziel der Benennaufgaben besteht darin, den Namen der zu identifizierenden Person zu nennen (Stimm- / Sprecheridentifizierung). Das individuelle Maß der Erkennungsleistung wird dabei meist in der Anzahl der korrekten Identifizierungen angegeben (*percentage correct [pc]*).

Unter **Matching-Task** werden jene Aufgaben verstanden, bei denen ein Teststimulus im Zusammenhang mit einem oder mehreren Distraktormuster(n) präsentiert wird. Der Teilnehmer hat dann z.B. die Aufgabe, das mit dem Teststimulus identische Muster auf einem Antwortbogen zu markieren. Dabei ist auch zu unterscheiden, ob der Teststimulus unmittelbar oder möglicherweise mehrere Tage oder Wochen vor der Matching-Aufgabe erfolgt. Eine Variante der Matching-Aufgabe ist das unmittelbare Vergleichen zweier Stimmuster. Dazu wird zu dem Teststimulus lediglich eine einzige Alternative dargeboten. Der Teilnehmer hat daraufhin eine gleich / ungleich-Entscheidung zu treffen.

Die in der vorliegenden Untersuchung gewählte Methode lässt sich nur schwer in die Taxonomie der bereits bestehenden Untersuchungen einordnen. Sie entspricht aber am ehesten einer 6-AFC-Diskriminierungsaufgabe mit einem *closed set*.

8.2. Hypothesen

Aufgrund der in Teil I dieser Arbeit vorgestellten theoretischen Überlegungen lassen sich folgende Hypothesen formulieren.

1. Aufgrund der zerebralen Läsionen und der damit einhergehenden neuropsychologischen Begleitsymptomatik sollten die Patienten der Gruppen LBD und RBD im Vergleich zu den zerebral gesunden Kontrollprobanden eine beeinträchtigte Leistung im Bereich der Stimmerkennung zeigen.
2. Die aktuellen Modelle der zerebralen Hemisphärenspezialisierung verbinden die Fähigkeiten der Mustererkennung mit den Funktionen der rechten Hemisphäre, während die analytischen Prozesse in der linken Hemisphäre spezialisiert zu sein scheinen. Diese theoretischen Modellüberlegungen decken sich mit klinischen Studien zur visuellen Mustererkennung, bei denen übereinstimmend die Patienten mit einer rechtsseitigen Hirnläsion beeinträchtigte Leistungen zeigen. Da nun das Erkennen einer vertrauten Stimme in erster Linie einer Mustererkennung gleicht, bei der die holistische Gestalt einer Stimme mit dem gespeicherten

8. Methodik

(Stimm-)Muster – hier: dem *unique pattern* (vgl. Kapitel 7.3.2) – einer Person verglichen wird, ist davon auszugehen, dass die Patienten mit einer rechtsseitigen zerebralen Läsion denen mit einer linksseitigen Läsion bzw. denen ohne zerebraler Läsion im Leistungsbereich der Stimmerkennung unterlegen sind.

3. Aufgrund der reduzierten Fähigkeit im Bereich der Mustererkennung lässt sich die Hypothese ableiten, dass die Gruppe der RBD-Patienten auch im Bereich des Erkennens visueller Muster den Patienten der Gruppe LBD und den HC-Probanden unterlegen ist.
4. Im Rahmen des Wahrnehmungsexperimentes finden Stimuli mit unterschiedlicher Länge Verwendung. Dabei ist nahe liegend, dass ein Stimulus von zeitlich größerer Dauer einen entsprechend umfangreicheren Ausschnitt der individuellen Stimmcharakteristik zeigt als ein zeitlich kürzerer Stimulus (z.B. Satz vs. Vokal). Aus diesem Grund sollte die Wiedererkennungseistung der Patienten und Probanden mit der zeitlichen Dauer der Stimuli zunehmen.

8.3. Testdesign I: Vortestungen

Um ausschließen zu können, dass die Leistungsmessungen durch begleitende (neuro-psychologische) Störungen (wie z.B. Konzentrationsstörungen) oder bestimmte Umweltbedingungen (z.B. unzureichende Vertrautheit mit der Zielstimme) verfälscht würden, wurden mit allen teilnehmenden Personen spezifische Vortests durchgeführt, die in den Kapiteln 8.3.1 bis 8.3.5 näher vorgestellt werden. An dieser Stelle sei darauf hingewiesen, dass sämtliche Untersuchungen unter Berücksichtigung der *Deklaration des Weltärztebundes von Helsinki* zu den ethischen Grundsätzen der medizinischen Forschung an Menschen in seiner Fassung von 2004 erfolgten.

Zu den Auswahlkriterien der beiden **Patientengruppen** gehörten: Z.n. unilateraler, postakuter Läsion der linken *oder* rechten Hemisphäre (Broca- / Wernicke-Region bzw. homologe Areale der rechten Hemisphäre), Z.n. Erst-Infarkt, Deutsch-Muttersprachler, Rechtshändigkeit und eine hohe Vertrautheit mit der Zielstimme (siehe unten).

Zu den Ausschlusskriterien zählten: Z.n. bilateralen Läsionen, Zustand nach (Z.n.) Re-Infarkt, Z.n. Blutungen, Tumoren oder anderen Raumfordernden Prozessen, Störung der Aufmerksamkeit, der Konzentration, des auditiven Gedächtnisses oder des Hörens sowie eine demenzielle und / oder depressive Symptomatik.

Die nachfolgende Tabelle 8.1 gibt eine Übersicht über die Teilnahme- (+) bzw.

8. Methodik

Ausschlusskriterien (–) für die unterschiedlichen Teilnehmergruppen (Patient / HC).

| Kriterien | Patient | HC | Quelle |
|---------------------------------|---------|----|----------|
| unilaterale Läsion | + | – | |
| Z.n. <i>ischäm.</i> Hirninfarkt | + | – | |
| Muttersprache Deutsch | + | + | Anamnese |
| Rechtshändigkeit | + | + | |
| Vertrautheit mit Zielstimme | + | + | |
| Störung des Hörens | – | – | |
| Störung der Konzentration | – | – | AKT |
| V.a. Demenz | – | – | |
| Störung des aud. Gedächtnisses | – | – | WMS–R |
| V.a. Depression | – | – | HADS–D |

Tabelle 8.1.: Übersicht über Auswahl- und Ausschlusskriterien

Um einen homogenen dialektalen Hintergrund zu gewährleisten, sollten die teilnehmenden Angehörigen (Zielstimmen) ebenso wie die Sprecher der Distraktorstimmen in der Region Ostwestfalen geboren und / oder aufgewachsen sein. Darüber hinaus musste die Vertrautheit mit der Zielstimme gewährleistet sein. Aus diesem Grunde wurden lediglich jene Versuchsteilnehmer in der Auswertung berücksichtigt, die seit mindestens zehn Jahren intensive soziale Kontakte (in Form von Ehe oder Lebenspartnerschaft) zu dem Sprecher der Zielstimme unterhielten.

8.3.1. Anamnesegespräch

Mit allen Patienten und Probanden wurde ein Anamnesegespräch geführt, in dessen Verlauf die relevanten Angaben zur **Person** (Name, Vorname, Geschlecht, Geburtsdatum und -ort, Familienstand, Schulabschluss, Dauer der Schulbildung, erlernter / ausgeübter Beruf, Händigkeit, Mutter- und Fremdsprachen), zu **medizinischen / therapeutischen Maßnahmen** (stationärer Behandlungszeitraum / Anschlussheilbehandlung (AHB), ambulante / stationäre logopädische Versorgung) und zur **Zielstimme** (Verwandtschaftsgrad mit dem Teilnehmer, Geburtsort, Dauer der Ehe / Partnerschaft) erhoben wurden. Darüber hinaus bot der Erfassungsbogen (vgl. Abbildung 13.3 auf Seite 222 Platz für die Informationen zu den **medizinischen Befunden** (Läsionsort, Lokalisation, Erstinfarkt, Datum des Insultereignisses, Ätiologie, Symptomatologie), die den Gesprächen mit den behandelnden Ärzten bzw. den Patientenakten entnommen wurden. Auch die Ergebnisse der Voruntersuchun-

8. Methodik

gen (vgl. Kapitel 8.3.2 bis 8.3.5) wurden auf ihm vermerkt.

Da sich die teilnehmenden Patienten zum Untersuchungszeitpunkt bereits im Rahmen ihrer AHB in Reha-Kliniken aufhielten, lagen nur sehr vereinzelt CT-Aufnahmen vor. Die erstbetreuenden Akutkrankenhäuser zeigten sich – trotz vorliegender Einverständniserklärungen – nicht bereit, ein Sichten der entsprechenden Scans zu ermöglichen.

8.3.2. Aachener Aphasie Test (AAT)

Um eine zuverlässige Syndromzuweisung gewährleisten zu können, wurde mit allen LBD-Patienten der Aachener Aphasie Test (Huber et al. 1983) durchgeführt. Der AAT erlaubt die Diagnostik aphasischer Störungen als Folge einer erworbenen Hirnschädigung und wurde für den deutschen Sprachraum entwickelt. Er überprüft die Ebene der Spontansprache mit Hilfe eines halbstandardisierten Interviews auf sechs neurolinguistisch definierten Ebenen. Die weiteren Subtests geben Informationen über etwaige Störungen in den Bereichen Nachsprechen, Lesen, Schreiben, Benennen und Sprachverständnis. Der AAT erlaubt u.a. die Auslese aphasischer Patienten aus einer Population nicht-aphasischer Patienten, ordnet die Betroffenen gemäß des neoklassischen Ansatzes den vier Standardsyndromen (Amnestische, Wernicke-, Broca- und Globale Aphasie) zu oder identifiziert Nicht-Standardsyndrome und modalitätsspezifische Sprachstörungen und beschreibt die sprachlichen Leistungen auf den unterschiedlichen linguistischen Ebenen. Da sich die Teilnehmer der Patientengruppe (vgl. Kapitel 9.1.1) zum Untersuchungszeitpunkt in Reha-Kliniken aufhielten, konnte in sämtlichen Fällen auf die AAT-Diagnostik der dort tätigen Sprachtherapeuten zurück gegriffen werden.

8.3.3. Wechsler-Gedächtnistest – Revised (WMS-R)

Um die Wahrnehmungsaufgabe korrekt durchführen zu können, muss der Patient zunächst die Stimmuster anhören, kurzzeitig abspeichern und im Anschluss eine bekannt / unbekannt-Entscheidung treffen. Dazu bedarf es eines ungestörten auditiven Gedächtnisses, das mit Hilfe eines Subtest der WMS-R (Härting et al. 2000) überprüft wurde. Die WMS-R ist die deutsche Version der Wechsler-Memory-Scale (Wechsler 1987) und bewertet über mnestiche Aufgaben fünf Leistungsaspekte des Patienten: das Verbale Gedächtnis, das Visuelle Gedächtnis – die zusammen den Allgemeinen Gedächtnis-Index ergeben –, die Aufmerksamkeit / Konzentration und die Verzögerte Wiedergabe. Der auf diese Weise ermittelte Gedächtnisquotient erlaubt einen Vergleich mit dem Intelligenzquotienten und gibt Hinweise auf das Bestehen

8. Methodik

eines mnestischen Syndroms. Die Reliabilitäten der einzelnen ermittelbaren Indizes rangieren zwischen $r = .70$ und $r = .90$ (vgl. Härtling et al. 2000).

Im Rahmen der vorliegenden Untersuchung wurden alle Versuchspersonen mit dem Subtest „Verbale Paarerkennung I“ überprüft. Dazu wurden den Teilnehmern acht Wortpaare vorgelesen, von denen sie sich jeweils das zweite Wort merken sollten. Im Anschluss wurde das jeweils erste Wort vorgegeben und die Teilnehmer sollten das entsprechende zweite Wort erinnern¹. Dazu wurden insgesamt mindestens drei Durchgänge vorgenommen, bei denen maximal 24 Punkte (3x8) erreicht werden konnten. Die Liste mit den verwendeten Wortpaaren findet sich im Anhang (vgl. Abbildung 13.4 auf Seite 223). Teilnehmer, die nicht in der Lage waren, den dritten Durchgang fehlerfrei zu absolvieren, wurden aus der statistischen Bewertung ausgeschlossen.

8.3.4. Alters-Konzentrations-Test (AKT)

Um ein Defizit im Bereich der Konzentrationsfähigkeit und / oder Vigilanz bzw. eine Aussage über das etwaige Vorliegen einer demenziellen Symptomatik treffen zu können, wurde mit allen Teilnehmern der AKT (Gatterer 1990) durchgeführt. Dabei handelt es sich um ein psychometrisches Verfahren, das eine Messung der Konzentrationsfähigkeit und Vigilanz speziell von älteren Menschen erlaubt. Die Aufgabe für den Teilnehmer besteht darin, eine auf dem Testbogen abgebildete Figur aus einer Auswahl ähnlicher Figuren herauszusuchen und zu markieren (vgl. Abbildungen 13.5 bis 13.9 ab Seite 224). Als Rohwerte dienen die benötigte Zeit, die Anzahl korrekt markierter Figuren sowie Anzahl und Art der Fehler. Sollte bei dem Teilnehmer eine demenzielle Symptomatik vorliegen, so gibt der AKT Auskunft über den Grad des zerebralen Abbaus.

8.3.5. Hospital Anxiety and Depression Scale – Dt. Version (HADS-D)

Die HADS-D (Herrmann-Lingen, Buss & Snaith 1995) ist die deutsche Version der englischsprachigen HADS (Zigmond & Snaith 1983) und wurde als Screeningverfahren zur Selbstbeurteilung für die Bereiche Angst und Depression konzipiert. Beide Subtests bestehen aus Fragebögen mit je sieben Testitems, die auf einer vierstufigen Skala beurteilt werden sollen (vgl. Abbildung 13.10 und 13.11 ab Seite 229). Um das Vorliegen einer depressiven Symptomatik erfassen zu können, wurde der Subtest Depression mit allen Teilnehmern durchgeführt.

¹Z.B. „Metall - Eisen.“, „Metall?“ — „Eisen!“.

8.3.6. Durchführung I

Die Akquise der hirnerkrankten Patienten erfolgte über bestehende Kooperationen mit unterschiedlichen Reha-Kliniken, wobei der weitaus größte Teil aus der *Klinik am Rosengarten*, der *Klinik am Osterbach* und der *Maternus-Klinik* (alle Bad Oeynhausen) stammte. Der Kontakt wurde über die dort tätigen Sprachtherapeuten bzw. Neuropsychologen hergestellt. Die Probanden der Gruppe HC stammten aus dem Bekannten- und Verwandtenkreis des Untersuchers bzw. seiner Kollegen der Abteilung Klinische Linguistik der Fakultät für Linguistik und Literaturwissenschaft, Universität Bielefeld.

Im Rahmen eines Aufklärungsgesprächs wurden die Probanden, die Patienten und deren Angehörige über den Ablauf und den zeitlichen Rahmen der Untersuchung informiert. Zusätzlich erhielten Sie ein Informationsblatt mit den entsprechenden Rahmenbedingungen (siehe Abbildung 13.2 auf Seite 221). Alle Teilnehmer unterzeichneten eine schriftliche Einverständniserklärung, mit der sie der Speicherung personenbezogener Daten und der anonymisierten, wissenschaftlichen Auswertung der Messergebnisse zustimmten.

Die zu erkennenden Zielstimmen der Ehe- / Lebenspartner wurden mit Hilfe des Digital Audio Tape-corder (DAT) der Firma Sony, Japan, mit einer Samplingfrequenz von 48 kHz aufgezeichnet und elektronisch gespeichert. Als Mikrofonmodul diente das Modell ME 64 und als Speiseadapter das Modell K6 der Firma Sennheiser. Die technischen Daten sind in Tabelle 14.1 und 14.2 im Anhang ab Seite 235 zusammengefasst.

Im Verlauf eines ersten Treffens wurden die Teilnehmer zunächst über den Ablauf der Untersuchung informiert. Zur Vorbereitung des Gesprächs wurden die Patientenakten gesichtet und die anamnestischen Daten erhoben. Waren sie unvollständig, so konnten sie im Gespräch mit den Patienten / Probanden oder den behandelnden Ärzten entsprechend ergänzt werden. Nach dem Unterzeichnen der Einverständniserklärung folgten die in Kapitel 8.3.2 bis 8.3.5 vorgestellten Vortestungen. Dabei wurde zunächst der Subtest Verbale Paarerkenntnis I der WMS-R (Härting et al. 2000), danach der AKT (Gatterer 1990) und zum Abschluss der Subtest Depression der HADS-D (Herrmann-Lingen, Buss & Snaith 1995) durchgeführt. Ohne Beisein des Patienten / Probanden erfolgte im Anschluss die Aufzeichnung der erforderlichen Stimuli mit dem Ehe- / Lebenspartner des Teilnehmers. Mit Hilfe der Software Praat (siehe Kapitel 14.2 auf Seite 236) wurden die Stimuli im Anschluss durch den Untersucher aus den aufgezeichneten Sätzen extrahiert und separat gespeichert (siehe Kapitel 8.4.2).

8.4. Testdesign II: Wahrnehmungsexperiment

Im Rahmen des Wahrnehmungsexperimentes wurden nun die Leistungen im Bereich der Stimmerkennung bei den hirnverletzten Patienten sowie bei den HC-Probanden untersucht.

8.4.1. Verwendete Distraktorstimmen

Da die Versuchsperson die Zielstimme unter verschiedenen Distraktorstimmen herausfinden musste, wurden dazu unterschiedliche weibliche und männliche Distraktoren benötigt.

Beschreibung der Distraktorstimmen

Je nach Geschlecht der zu erkennenden Zielstimme wurde das Set für weibliche bzw. männliche Zielstimmen verwendet. Beide Sets bestanden aus jeweils fünf Distraktorstimmen (D1 bis D5). D1, D2 und D3 beider Sets stammten von Personen, die dem ungefähren Alter der zu erwartenden Zielstimmen entsprachen (Senioren), D4 wich deutlich davon ab (Studierende). D5 stammte in beiden Sets von altersadäquaten aber gegengeschlechtlichen Personen.

Das Set bei einer **weiblichen Zielstimme** (Ehe- / Lebenspartnerin) setzte sich also wie in Tabelle 8.2 dargestellt zusammen:

| Distraktor | Beschreibung | Quelle |
|------------|--------------|--------------------------|
| wD1 | Seniorin | Bielefelder Seniorenheim |
| wD2 | Seniorin | |
| wD3 | Seniorin | |
| wD5 | Senior | |
| wD4 | Studentin | |

Tabelle 8.2.: Set bei weiblichen Zielstimmen

Das Set bei einer **männlichen Zielstimme** (Ehe- / Lebenspartner) setzte sich analog dazu wie folgt zusammen (vgl. Tabelle 8.3):

Ausgewählte akustische Parameter (Pitch, Jitter und Shimmer) der eingesetzten Distraktorstimmen sind im Anhang aufgeführt (vgl. Tabelle 12.5 auf Seite 194 im Anhang). Sie stellen die durchschnittlichen Werte dar, die durch die akustische Analyse der beiden Testitems und der beiden Sätze S1 und S2 gewonnen wurden.

8. Methodik

| Distraktor | Beschreibung | Quelle |
|------------|--------------|--------------------------|
| mD1 | Senior | Bielefelder Seniorenheim |
| mD2 | Senior | |
| mD3 | Senior | |
| mD5 | Seniorin | |
| mD4 | Student | Universität Bielefeld |

Tabelle 8.3.: Set bei männlichen Zielstimmen

Sammlung der Distraktorstimmen

Die Distraktorstimmen stammten aus mehreren Seniorenheimen der Stadt Bielefeld (wD1 bis D3 und wD5 bzw. mD1 bis mD3 und mD5) bzw. aus der Studierendenschaft der Universität Bielefeld, Fakultät für Linguistik und Literaturwissenschaft (wD4 bzw. mD4, siehe oben). Die Teilnehmer wurden zunächst schriftlich und am Tag der Aufzeichnung nochmals mündlich über den Ablauf der Aufzeichnungsprozedur unterrichtet (siehe Abbildung 13.1 auf Seite 220). Die Aufnahmen erfolgten einzeln in den Räumen der Seniorenheime bzw. der Universität Bielefeld.

Um homogene dialektale Rahmenbedingungen zu schaffen, sollten die Distraktoren im ostwestfälischen Sprachraum geboren und / oder aufgewachsen sein. Darüber hinaus sollten Sprache, Sprechen und Stimme befundfrei sein, um zu deutliche idiolektale Abweichungen von den Zielstimmen auszuschließen.

Die Distraktorstimmen und zu einem späteren Zeitpunkt auch die jeweiligen Zielstimmen wurden mit Hilfe des DAT-Rekorders Walkman TCD-D100 der Firma Sony, Japan, mit einer Samplingfrequenz von 48 kHz digital aufgezeichnet und elektronisch gespeichert. Als Mikrofonmodul wurde das Modell ME 64 und als Speiseadapter das Modell K6 verwendet. Beides sind Produkte der Firma Sennheiser, Deutschland. Die Wiedergabe der Stimuli erfolgte über handelsübliche PC-Lautsprecher der Firma WaveMaster, Großbritannien. Die technischen Angaben zu den verwendeten elektronischen Geräten sind im Anhang (siehe Kapitel 14 ab Seite 234) aufgeführt.

8.4.2. Beschreibung der Textstimuli

Sowohl die Distraktoren als auch die teilnehmenden Ehe- / Lebenspartner wurden gebeten, die Inhalte mehrerer Texttafeln vorzulesen. Zu diesem Zweck wurden zehn deutschsprachige Sätze auf zehn Texttafeln gedruckt (DIN A4, Querformat, zentriert, Times New Roman, 72pt). Die ersten beiden Tafeln („Der Unfall ereignete sich an der Kreuzung.“ und „Die Lehrerin ermahnt die beiden Schüler.“) dienten dabei als

8. Methodik

Testitems und wurden nicht weiter verwendet. Aus den übrigen acht Sätzen wurden die jeweiligen Untersuchungsstimuli extrahiert.

Die Stimuli V1 und V2 bestanden aus den **Vokoiden** [a:] und [o:], die aus den Originalsätzen „Ich sagte 'All', nicht 'Aal'!“ bzw. „Ich sagte 'offen', nicht 'Ofen'!“ entnommen wurden. V1 und V2 hatten jeweils eine Länge von ca. 140 ms. Zeigten die Vokoide im Originalsatz eine größere Dauer, so wurden sie im Rahmen der Stimmmusterpräsentation auf 140 ms gekürzt. Wiesen sie eine geringere Dauer auf, so wurde die Aufnahme wiederholt. Die **CVC-Cluster** 1 [ˈtan] und 2 [pu:k] wurden aus den Sätzen „Am Wochenende besuchte er seine Tante.“ bzw. „Um Mitternacht begann der gruselige Spuk.“ extrahiert. CVC1 hatte eine Länge von durchschnittlich 270 ms, CVC 2 von ca. 390 ms. Die Stimuli W1 [ˈɔyro:] und W2 [prɛˈzi:dɛnt] bestanden aus kompletten **Wörtern**, die durch die Originalsätze ‘In Deutschland bezahlt man mit dem Euro.‘ bzw. „In der entführten Maschine sitzt der Präsident.“ gewonnen wurden. Die zeitliche Dauer von W1 betrug rd. 500 ms, von W2 rd. 750 ms. Eine Angleichung der Länge wurde nicht vorgenommen. Die beiden letzten Stimuli S1 und S2 bestanden aus jeweils einem der beiden **Sätze** „Die Frau führt das Unternehmen erfolgreich.“ [di: frau fyrt das ʊntəˈne:mən ɛɣˈfɔlkraɪç] bzw. „Über drei Millionen Menschen leben in der Hauptstadt Berlin.“ [ˈy:bɛ drai mɪˈljɔ:n ˈmɛnʃn ˈle:bɪn ɪn dɛrɣ hauptʃtat bɛrˈli:n]. Stimulus S1 hatte eine Länge von ca. 2600 ms, S2 von > 3000 ms. Auch bei diesen Stimuli wurde keine Angleichung der Länge vorgenommen (vgl. auch Tabelle 8.4).

| Item | Stimulus | IPA | Dauer (ms) |
|------|----------|-----------------------------------------------------------------|------------|
| 1 | V1 | [a:] | ca. 140 |
| 2 | V2 | [o:] | |
| 3 | CVC1 | [ˈtan] | ca. 270 |
| 4 | CVC2 | [pu:k] | ca. 390 |
| 5 | W1 | [ˈɔyro:] | ca. 500 |
| 6 | W2 | [prɛˈzi:dɛnt] | ca. 750 |
| 7 | S1 | [di: frau fyrt das ʊntəˈne:mən ɛɣˈfɔlkraɪç] | ca. 2600 |
| 8 | S2 | [ˈy:bɛ drai mɪˈljɔ:n ˈmɛnʃn ˈle:bɪn ɪn dɛrɣ hauptʃtat bɛrˈli:n] | > 3000 |

Tabelle 8.4.: Übersicht über die verwendeten Stimuli nach Stimulikategorie

8.4.3. Stichprobe

Auf Grundlage der Ergebnisse aus den Vortestungen wurden die Versuchspersonen in drei Gruppen eingeteilt:

Gruppe 1: aphasische Patienten mit einer Läsion der linken Hemisphäre (LBD),

Gruppe 2: nicht-aphasische Patienten mit einer Läsion der rechten Hemisphäre (RBD) und

Gruppe 3: zerebral gesunde Kontrollpersonen (HC).

8.4.4. Durchführung II

Um die physische Belastung insbesondere der teilnehmenden Patienten möglichst gering zu halten, wurden pro Aufgabe neben der Zielstimme fünf Distraktorstimmen dargeboten (6-AFC). Die Leistung des Teilnehmers bestand darin, unter diesen sechs Stimmustern die Stimme des Ehe- / Lebenspartners zu erkennen. Dabei war allen teilnehmenden Patienten und Probanden bereits vor Beginn der Erkennungsaufgabe bekannt, dass die zu erkennende Stimme die des Ehe- / Lebenspartners sein würde. Dies ließ sich nicht vermeiden, da der Kontakt zu den Ehe- / Lebenspartnern in den meisten Fällen direkt über die Patienten / Probanden erfolgte. Darüber hinaus musste davon ausgegangen werden, dass die beteiligten Angehörigen die Patienten / Probanden über das Aufzeichnen ihrer (Ziel-) Stimme unterrichten würden.

Die Erkennungsaufgabe gliederte sich in acht Testitems, die sich aus vier Stimulustypen zusammensetzten und sich vor allem in ihrer zeitlichen Länge und dem enthaltenen Phoneminventar unterschieden (vgl. Kapitel 8.4.2).

Je nach Zustand und Belastbarkeit des Patienten erfolgte nach den bereits an anderer Stelle beschriebenen Vortestungen der zweite Teil der Untersuchung. Das Wahrnehmungsexperiment wurde entweder direkt im Anschluss oder zeitnah an einem anderen Tag durchgeführt. Dazu wurde dem Teilnehmer nochmals der genaue Ablauf erläutert. Sobald der Teilnehmer seine Bereitschaft signalisierte, wurde mit der Präsentation der Stimuli des Items V1 begonnen (es folgten V2, CVC1, CVC2, W1, W2, S1 und S2). Jedes Item wurde von insgesamt sechs Sprechern produziert. Ein Sprecher war dem Teilnehmer bekannt (Ehe- / Lebenspartner), die fünf anderen Sprecher (Distraktorstimmen) waren ihm unbekannt. Je nach Geschlecht des Teilnehmers wurde das Set der weiblichen bzw. männlichen Distraktorstimmen gewählt (wD1 - 5 bzw. mD1 - 5). Jedem der sechs Stimmuster wurde eine Ziffer (1 bis 6) in Form eines vertikal angeordneten Zahlenstranges zugeordnet. Nach dem Hören aller

8. Methodik

sechs Stimmuster wurde der Teilnehmer aufgefordert, die Ziffer zu nennen oder zu zeigen, die nach seiner Meinung zu der ihm bekannten Stimme gehörte. War die Versuchsperson unsicher, so bestand die Möglichkeit, die Stimmen ein zweites Mal zu hören. Das Ergebnis der Erkennungsleistung wurde von dem Untersucher protokolliert (siehe Abbildung 13.12 auf Seite 231) und in *sofortiges*, *verzögertes* bzw. *nicht Erkennen* differenziert. Wurde eine falsche Ziffer genannt, so wurde zusätzlich die vom Teilnehmer fälschlicherweise bestimmte Distraktorstimme notiert (D1 bis D5). Im Anschluss an jedes Testitem wurde der Teilnehmer gebeten, auf einer fünfstufigen Skala (-2 bis +2) die subjektive Sicherheit und den subjektiven Schweregrad der Aufgabe zu bestimmen. Eine Rückmeldung über die Korrektheit der gewählten Stimme erfolgte nicht.

9. Ergebnisse

9.1. Stichprobe

Insgesamt nahmen an der Untersuchung 49 Patienten und Probanden teil. Davon waren 17 weiblich und 32 männlich. Auf Grund der Ergebnisse der in Kapitel 8.3 beschriebenen Vortestungen, konnten 12 Teilnehmer (3 weiblich, 9 männlich) bei der statistischen Auswertung nicht berücksichtigt werden. Das Durchschnittsalter der verbliebenen 37 Teilnehmer (14 weiblich, 23 männlich) betrug 64,73 Jahre (49,6 - 81,4 Jahre, $SD = 7,55$). Um die Leistungen der Teilnehmer vergleichen zu können, wurden sie in die Gruppen LBD, RBD und HC eingeteilt (vgl. Kapitel 9.1.1 bis 9.1.3).

9.1.1. Gruppe 1: Aphasiker (LBD)

Gruppe 1 setzte sich aus elf Patienten (zwei weiblich, neun männlich) im Alter von durchschnittlich 66,26 Jahren (50,1 - 74,9 Jahre, $SD = 7,72$) zusammen. Alle Patienten litten aufgrund eines ischämischen Hirninfarktes in der linken Hemisphäre unter eine Aphasie, die mit Hilfe des AAT (Huber et al. 1983) diagnostiziert und klassifiziert wurde. Demnach zeigten je fünf Patienten eine Amnestische bzw. eine Wernicke-Aphasie und ein Patient eine Broca-Aphasie. Die durchschnittliche *post-onset-time* lag bei 3,1 Monaten (1 - 12 Monate, $SD = 3,56$) (vgl. Tabelle 12.1 auf Seite 190 im Anhang). Eine Analyse mit Hilfe des Kolmogorov-Smirnov-Tests zeigt, dass die Gruppe LBD hinsichtlich der Variablen *Alter* und *post-onset-time* als normalverteilt gelten kann ($p = .718$ bzw. $p = .14$).

9.1.2. Gruppe 2: Nicht-Aphasiker (RBD)

Gruppe 2 setzte sich aus insgesamt neun Patienten (drei weiblich, sechs männlich) mit einer Läsion der rechten Großhirnhemisphäre in den homologen Regionen des Wernicke- bzw. Broca-Areals zusammen. Die Information über den Lokalisationsort stammte bei sämtlichen Patienten der Gruppe 2 aus persönlichen Gesprächen mit den behandelnden Ärzten bzw. aus den Akten der Akut- oder Reha-Kliniken oder konnten, in Einzelfällen, durch Sichtung der CT-Aufnahmen gewonnen werden. Das durchschnittliche Alter der Patienten betrug zum Testzeitpunkt 64,11 Jahre (49,6 - 81,4 Jahre, $SD = 11,1$), bei einer *post-onset-time* von durchschnittlich 1,56 Monaten (1 - 4 Monate, $SD = 1,01$) (vgl. Tabelle 12.2 auf Seite 190 im Anhang). Auch diese Gruppe kann hinsichtlich der Variablen *Alter* und *post-onset-time* als normalverteilt gelten ($p = .999$ bzw. $p = .159$).

9. Ergebnisse

9.1.3. Gruppe 3: gesunde Kontrollprobanden (HC)

Zu der Gruppe der zerebral gesunden Kontrollprobanden gehörten neun weibliche und acht männliche Teilnehmer im Alter von durchschnittlich 64,1 Jahren (51,7 - 72,5 Jahre, $SD = 5,21$; normalverteilt nach Kolmogorov-Smirnov-Test, $p = .547$ (vgl. Tabelle 12.3 auf Seite 191 im Anhang)).

9.1.4. Zusammenfassung

Die Variablen Alter und *p.-o.-time* werden in Tabelle 9.1 nochmals überblicksweise aufgeführt und mit Hilfe der Box-Plots (siehe Abbildungen 9.1 und 9.2 auf der nächsten Seite) graphisch dargestellt. Eine Varianzanalyse mit Hilfe einer einfaktoriellen ANOVA zeigt keine signifikanten Unterschiede zwischen dem Alter und der *post-onset-time* der Patienten und Probanden (vgl. Abbildung 12.1 auf Seite 199 im Anhang).

| Gruppe | Alter | <i>SD</i> | <i>p.-o.-time</i> | <i>SD</i> |
|--------|-------|-----------|-------------------|-----------|
| LBD | 66,26 | 7,72 | 3,1 | 3,56 |
| RBD | 64,11 | 11,1 | 1,56 | 1,01 |
| HC | 64,1 | 5,21 | – | – |

Tabelle 9.1.: Übersicht über Alter und *p.-o.-time*

9. Ergebnisse

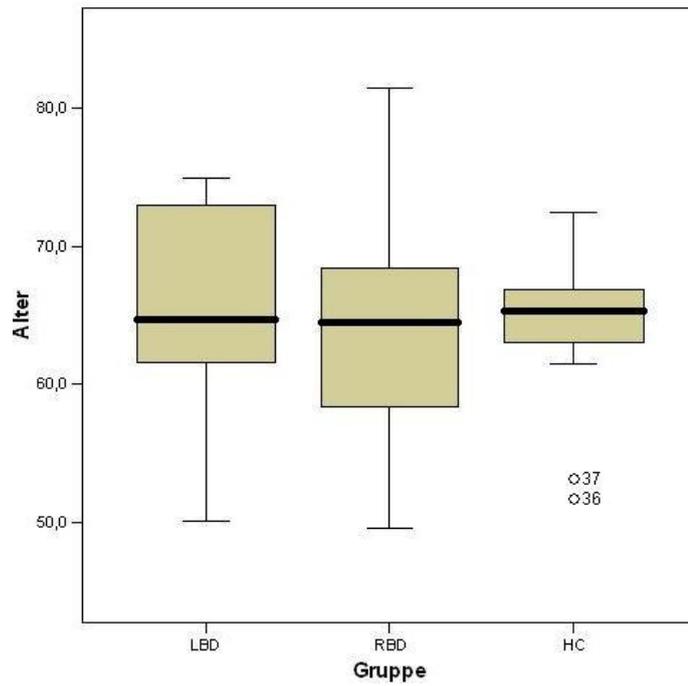


Abbildung 9.1.: Alterszusammensetzung der Gruppen (n.s.)

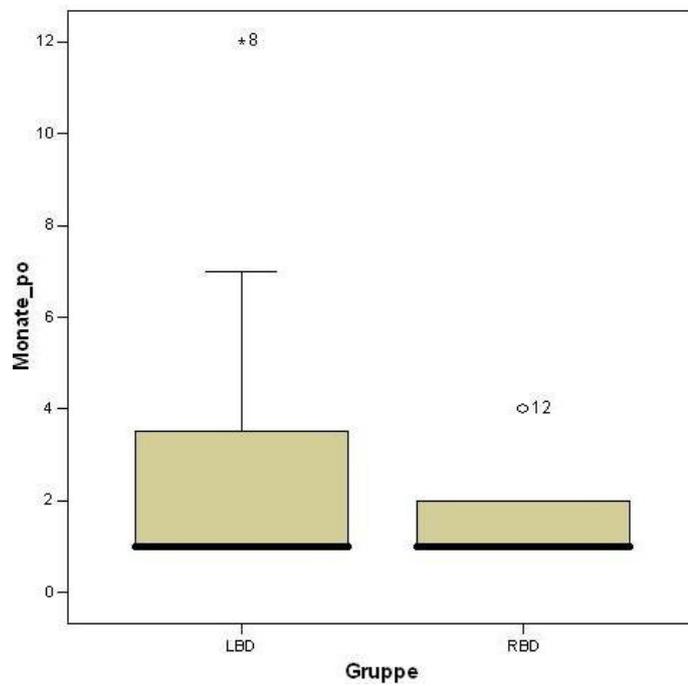


Abbildung 9.2.: p.o.-time der Patienten mit zerebraler Läsion (n.s.)

9.2. Zusammenfassung und Auswertung der Vortestergebnisse

In den beiden folgenden Unterkapiteln werden die Vortest-Ergebnisse der teilnehmenden Gruppen vorgestellt und ausgewertet.

9.2.1. Ergebnisse der WMS-R und der HADS-D

Im Rahmen der Voruntersuchungen erreichten die Patienten der LBD-Gruppe im Subtest Verbale Paarerennung I der WMS-R einen durchschnittlichen Punktrohwert von 19 (15 - 22, SD = 2,19), im HADS-D (Subtest Depression) von 3,27 (1 - 8, SD = 2,61).

Die Gruppe der RBD-Patienten erzielte bei der Verbalen Paarerennung I einen Punktrohwert von 20 (16 - 24, SD = 2,55) und von 3,33 (1 - 9, SD = 2,69) im Subtest der HADS-D.

In der WMS-R erreichten die HC-Probanden einen durchschnittlichen Punktrohwert von 19,76 (18 - 23, SD = 1,72), im Subtest der HADS-D von 2,24 (0 - 6, SD = 2,14).

Tabelle 9.2 fasst u.a. die Rohwerte (WMS-R und HADS-D) der Gruppen überblicksweise zusammen. Abbildungen 9.3 und 9.4 auf der nächsten Seite geben dies graphisch wieder. Eine Varianzanalyse mit Hilfe einer einfaktoriellen ANOVA zeigt bei den Ergebnissen in der WMS-R und der HADS-D keine statistischen Unterschiede (vgl. Abbildung 12.2 auf Seite 199 im Anhang).

| Gruppe | WMS-R | SD | HADS-D | SD |
|--------|-------|------|--------|------|
| LBD | 19 | 2,19 | 3,27 | 2,61 |
| RBD | 20 | 2,55 | 3,33 | 2,69 |
| HC | 19,76 | 1,72 | 2,24 | 2,14 |

Tabelle 9.2.: Ergebnisse in der WMS-R und dem HADS-D

9. Ergebnisse

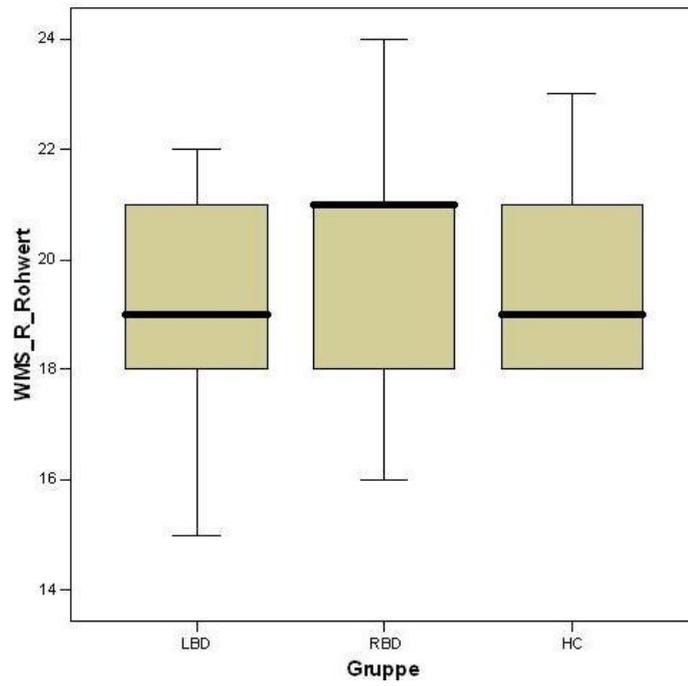


Abbildung 9.3.: Gruppenergebnisse in der WMS-R (n.s.)

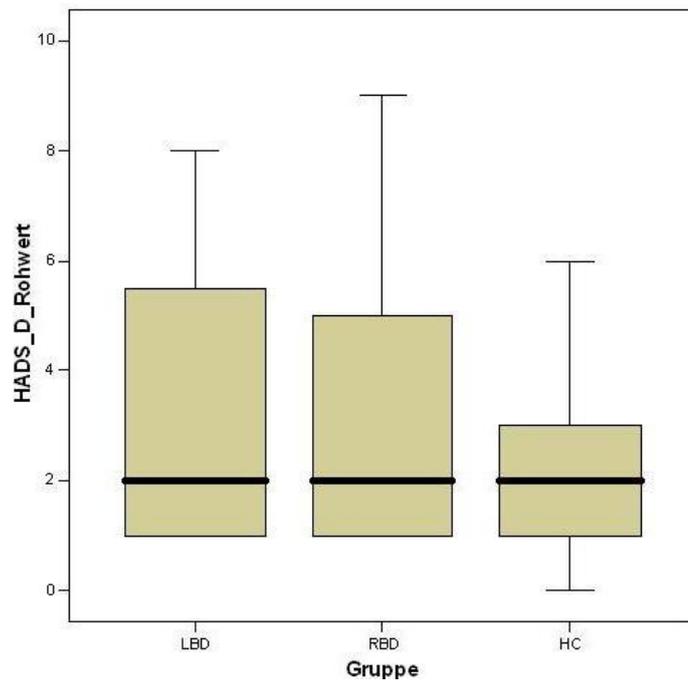


Abbildung 9.4.: Gruppenergebnisse in der HADS-D (n.s.)

9.2.2. Übersicht der Prozentränge des AKT

Da der AKT über Prozentrangangaben verfügt, werden diese im Folgenden für die Vergleiche der einzelnen Gruppen herangezogen. Dabei beziehen sich die Angaben der Gruppe LBD auf die AKT-Subgruppe „Aphasiker“ und die PR-Werte der Gruppe HC auf die AKT-Subgruppe „Rüstige“. Der Gruppe RBD wurden die Prozentränge der Subgruppe „Pflegeheimpopulation“ zugewiesen, da es sich bei diesem AKT-Eichkollektiv um Patienten handelt, deren Aufnahme in ein Pflegeheim aufgrund somatischer und / oder kognitiver Erkrankungen erfolgte und somit den Voraussetzungen der RBD-Gruppe am nächsten kommt (vgl. Gatterer 1990:20). Die im Rahmen der Testung gemessene Bearbeitungszeit **T** wird als Arbeitsgeschwindigkeit interpretiert. Die aktiven Probanden des AKT-Eichkollektivs benötigten für die Bearbeitung der Aufgabe weniger Zeit als inaktive. **R** steht für die Anzahl richtig erkannter Muster und bildet einen Teil des kognitiven Leistungsaspektes. Ein Indikator für die Leistungsgüte ist der Wert **F%**, für den zerebralen Abbau der Gesamtwert **G**, der den kognitiven Leistungsaspekt des AKT darstellt (vgl. Gatterer 1990:21).

Eine einfaktorielle ANOVA zeigt zwischen den Leistungen der einzelnen Gruppen einen signifikanten Unterschied im Wert **R**, der die Anzahl der korrekt erkannten optischen Muster angibt ($F(2,24) = 3,904$; $MS_{FEHL} = 876,185$; $p = .030$; vgl. Abbildung 12.3 auf Seite 199). Der Prozentrang für die Anzahl der richtig erkannten Muster (**R**) der RBD-Gruppe (73,54 %) unterscheidet sich im t-Test (nach Student für zwei unabhängige Stichproben) bei einer Irrtumswahrscheinlichkeit von 5 % signifikant von denen der beiden übrigen Gruppen LBD (90,53 %, $p = .024$) und HC (88,83 %, $p = .024$, vgl. Abbildung 9.5 auf der nächsten Seite).

| Gruppe | T | <i>SD</i> | R | <i>SD</i> | F% | <i>SD</i> | G | <i>SD</i> |
|--------|----------|-----------|----------|-----------|-----------|-----------|----------|-----------|
| LBD | 87,55 | 11,12 | 90,53 | 13,98 | 100,0 | 0 | 92,04 | 11,66 |
| RBD | 78,54 | 23,61 | 73,54* | 16,81 | 86,92 | 23,57 | 84,09 | 18,1 |
| HC | 89,18 | 12,51 | 88,83 | 14,61 | 97,01 | 12,32 | 91,67 | 11,15 |

Tabelle 9.3.: Prozentränge des AKT (sign.)

9. Ergebnisse

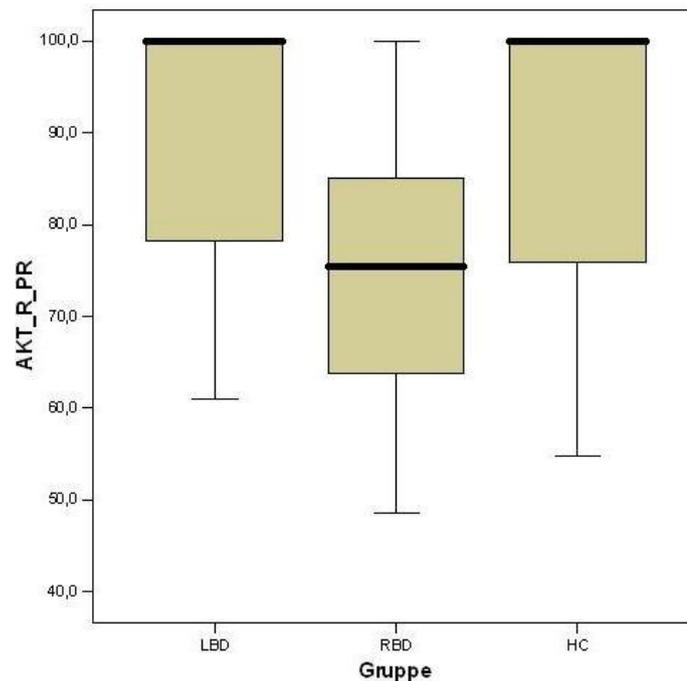


Abbildung 9.5.: Prozentrangmittelwerte R des AKT

9.3. Statistische Überprüfung der Hypothesen

Wie bereits in Kapitel 8.4.4 erläutert, bestand die Aufgabe der Versuchsteilnehmer darin, im Rahmen eines Wahrnehmungsexperimentes mit insgesamt acht Items die Stimme ihres Ehe- / Lebenspartners zu erkennen. Die Leistungen der Versuchspersonen wurden differenziert in sofortiges (*sE*), verzögertes (*vE*) oder nicht Erkennen (*nE*). Die Erkennungsaufgabe galt als erfüllt, wenn die Stimme im ersten (*sE*) oder zweiten Versuch (*vE*) korrekt erkannt wurde. Zeigte der Teilnehmer im ersten Versuch ein nicht Erkennen, so wurde ihm dies nicht mitgeteilt. Ein zweiter Versuch erfolgte lediglich dann, wenn der Teilnehmer von sich aus Unsicherheit signalisierte. Als Messgröße für die Erkennungsleistung der Versuchspersonen dient hier der prozentuale Anteil der korrekten Identifizierungen (*percentage correct, PC_passed*). Tabelle 9.4 fasst zunächst die Leistungen der Gruppen über die gesamten acht Erkennungsaufgaben zusammen (V, CVC, W und S) (siehe auch Abbildung 9.7 auf Seite 128).

9.3.1. Gesamterkennungsleistung

Die neuropsychologische Literatur berichtet von einer Fülle kognitiver und affektiver Funktionsstörungen bei Patienten mit einer zerebralen Läsion. Dazu zählen, neben den Sprech- und Sprachstörungen, auch Amnesien, Agnosien, Neglecte, Aufmerksamkeitsstörungen, Störungen der Planungs- und Kontrollfunktionen, Störungen des Antriebs und der Affektivität, Störungen des visuellen und auditiven Verarbeitens etc. Auch wenn durch die Vortestung der Patienten im Rahmen der vorliegenden Untersuchung einige neuropsychologische Begleitstörungen ausgeschlossen werden konnten (z.B. Störungen der Aufmerksamkeit, Depression etc., vgl. Kapitel 8.3 ab Seite 109), so ist dennoch davon auszugehen, dass die zerebral beeinträchtigten Patienten aufgrund ihrer Vorerkrankung(en) und der daraus resultierenden (neuropsychologischen) Begleitsymptomatik bei der Durchführung des Wahrnehmungsexperimentes größere Schwierigkeiten zeigen sollten als die zerebral *gesunden* Kontrollpersonen.

Um dies zu überprüfen, wurde die einfaktorielle ANOVA angewendet. Sie zeigt, dass sich die Leistungen zwischen BD-Patienten auf der einen und den HC-Probanden auf der anderen Seite nicht unterscheiden ($F[1,35] = 0,082$, $p = .776$, vgl. Abbildung 9.6).

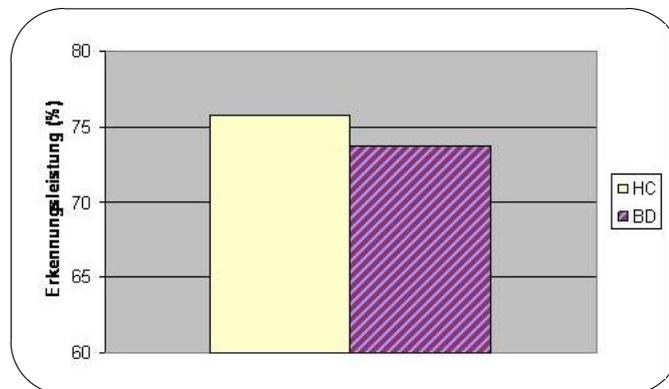


Abbildung 9.6.: Gesamterkennungsleistung der Gruppen HC und BD

Statistische und numerische Leistungsunterschiede zwischen Kontrollpersonen und BD-Patienten zeigen sich erst dann, wenn man die Gruppe der BD-Patienten differenziert betrachtet.

Eine Analyse der Gesamtleistung über alle Stimmerkennungsaufgaben mit Hilfe der einfaktoriellen ANOVA zeigt einen signifikanten Unterschied zwischen den Gruppenleistungen der Gruppen LBD, RBD und HC ($F [2,34] = 3,516$; $MS_{FEHL} =$

9. Ergebnisse

1325,085; $p = .041$; vgl. Abbildung 12.4 auf Seite 200). Zum Vergleich der Leistungen der unterschiedlichen Subgruppen miteinander, wurde als post-hoc-Test der U-Test nach Mann & Whitney angewandt. Er zeigt, dass sich der Prozentwert für die Anzahl der korrekt erkannten Stimmuster ($sE + vE$) der RBD-Gruppe (61,11 %) signifikant von den Leistungen der LBD-Gruppe (84,09 %) unterscheidet ($p = .018$). Im Vergleich mit den Leistungen der HC-Gruppe (75,74 %) zeigt sich eine statistische Tendenz ($p = .093$).

| Gruppe | N | sE (%) | vE (%) | nE (%) | sE + vE (%) |
|--------|----|--------|--------|--------------------|---------------------|
| LBD | 11 | 77,27 | 6,82 | 15,91 | 84,09 |
| RBD | 9 | 54,17* | 6,94 | 38,89 ^t | 61,11* ^t |
| HC | 17 | 65,44 | 10,29 | 24,27 | 75,74 |

sE = sofortiges Erk., vE = verzögertes Erk., nE = nicht Erk.

Tabelle 9.4.: Übersicht über die Gesamtleistung

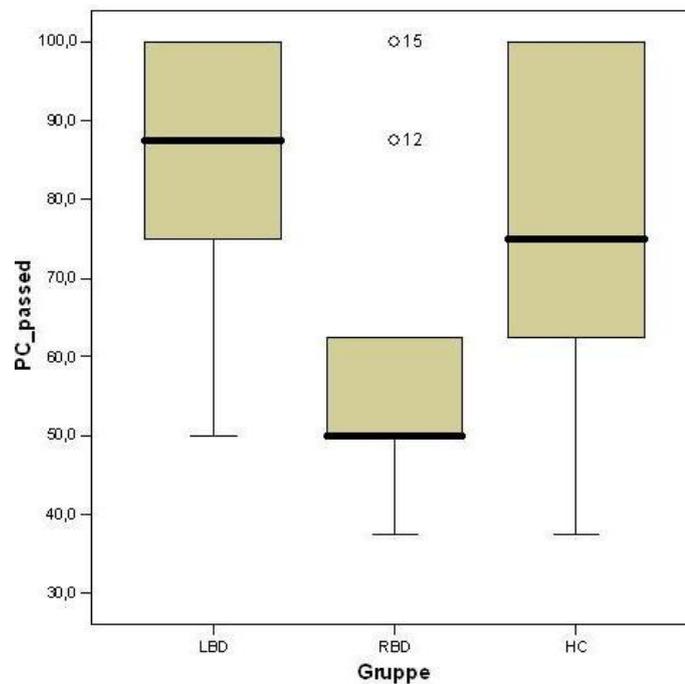


Abbildung 9.7.: Gruppenleistungen Gesamt ($sE + vE$)

9. Ergebnisse

Betrachtet man die sofortige Erkennungsleistung (sE) isoliert, so lässt sich im U-Test nach Mann & Whitney ein signifikanter Unterschied zwischen den sE -Leistungen der LBD- und der RBD-Gruppe (77,27 % bzw. 54,17 %) nachweisen ($p = .016$). Ein Vergleich zwischen der HC- und der RBD-Gruppe bei der Variable nE zeigt im U-Test eine statistische Tendenz (24,27 % bzw. 38,89 %; $p = .093$, siehe oben). Unterschiede zwischen den Gruppen in Bezug auf die Variable vE lassen sich nicht nachweisen. Ebenso zeigen sich zwischen den Gruppen LBD und HC bei keiner der o.g. Variablen statistische Unterschiede (vgl. Abbildungen 12.5 bis 12.10 ab Seite 200 im Anhang).

Wie in Abbildung 9.1 auf Seite 122 dargestellt, liegen die HC-Probanden 36 und 37 um eine Standardabweichung unterhalb des Durchschnittsalters der Gruppe HC. Aus diesem Grund soll nun überprüft werden, ob dies auf die Leistung im Bereich der Stimmerkennung Einfluss nimmt. **Proband 36** (w, 51;7 Jahre) zeigt bei der Stimmerkennungsaufgabe einen PC-Wert von 100 %. Dieser Wert unterscheidet sich hochsignifikant von dem Mittelwert der Gesamtgruppe und dem Mittelwert der HC-Gruppe ($p < .001$). **Proband 37** (m, 53;1 Jahre) erreicht im Rahmen der Stimmerkennungsaufgabe einen PC-Wert von 62,5 %. Dieser Wert unterscheidet sich sowohl von dem Mittelwert der Gesamtgruppe ($p = .001$), als auch von dem der HC-Gruppe ($p = .019$).

Der **LBD-Proband 8** liegt mit einer *post-onset-time* von zwölf Monaten um zwei Standardabweichungen über dem Mittelwert der LBD-Gruppe (3,1 Monate, vgl. Abbildung 9.2 auf Seite 122). Auch für diesen Patient soll überprüft werden, ob sich seine Leistungen im Bereich der Stimmerkennung von denen der zerebral geschädigten Gesamtgruppe bzw. von denen der LBD-Gruppe unterscheiden. Der Proband (w, 61;8 Jahre) erreichte bei der Stimmerkennungsaufgabe einen PC-Wert von 100 %. Dieser Wert unterscheidet sich von dem Mittelwert der zerebral beeinträchtigten Gesamtgruppe hochsignifikant ($p < .001$), von dem Mittelwert der LBD-Gruppe sehr signifikant ($p = .007$). Auch die *post-onset-time* des **RBD-Probanden 12** (4 Monate, w, 58;3 Jahre) liegt um eine Standardabweichung über der durchschnittlichen *p.-o.-time* der RBD-Gruppe von 1,56 Monaten. Die Probandenleistung im Bereich der Stimmerkennung unterscheidet sich bei einem PC-Wert von 87,5 % sehr signifikant ($p = .008$) von dem Mittelwert der zerebral beeinträchtigten Patienten bzw. von dem der RBD-Gruppe ($p = .003$).

9. Ergebnisse

Besonders auffällig sind die Leistungen im Bereich der Stimmerkennung bei Patient 12 und 15, die jeweils um eine Standardabweichungen über dem Mittelwert der RBD-Gruppe liegen (48,6 % bzw. 55,1 %). Im Umkehrschluss zur weiter oben formulierten Hypothese soll an dieser Stelle angenommen werden, dass Patienten mit einer überdurchschnittlichen Leistung im Erkennen akustischer Muster auch beim Erkennen optischer Muster überlegen sein sollten. Dies lässt sich bei **Patient 12** allerdings nicht bestätigen. Hier ist das Gegenteil der Fall. Der Prozentrang von 48,6 unterscheidet sich von dem Mittelwert der Gesamtgruppe (PR = 85,66) hochsignifikant ($p < .001$). Dies gilt auch für den Vergleich mit der Gesamtgruppe der zerebral beeinträchtigten Patienten (Gruppe LBD + RBD; PR = 82,89; $p < .001$). Das gleiche Bild zeigt sich auch bei **Patient 15**, dessen Prozentrang „R“ im AKT bei 55,1 liegt und sich somit hochsignifikant ($p < .001$) von dem durchschnittlichen Prozentrang der Gesamtgruppe ($p < .001$) bzw. der zerebral beeinträchtigten Gruppe ($p < .001$) unterscheidet. Beide Patienten zeigen demnach überdurchschnittliche Leistungen im Bereich der auditiven Mustererkennung, während sie im Bereich der visuellen Mustererkennung hochsignifikant schlechter abschneiden als die übrigen Patienten / Probanden.

Weitere Ergebnisse

Differenziert man die Patienten der LBD-Gruppe nach ihren diagnostizierten Aphasiesyndromen (Wernicke- bzw. Amnestische Aphasie [jeweils $N = 5$] sowie Broca-Aphasie [$N = 1$]) und vergleicht die Mittelwerte ihrer Gesamterkennungsleistung mit Hilfe einer einfaktoriellem ANOVA, so zeigt sich kein statistisch signifikanter Unterschied ($F[2,8] = .258$, $p = .779$; vgl. Abbildung 12.11 auf Seite 203 im Anhang). Darüber hinaus zeigt sich über die Gesamtgruppe kein statistischer Zusammenhang zwischen der Erkennungsleistung und der Geschlechtszugehörigkeit ($F[1,35] = 1,216$, $p = .278$).

9.3.2. Erkennungsleistungen auf Vokal-, CVC-, Wort- und Satzebene nach Stimulikategorie

In diesem Unterkapitel werden die verwendeten Stimuli nicht nach ihrer zeitlichen Länge betrachtet, sondern zunächst ausschließlich nach ihrer Zugehörigkeit zu einer linguistischen Kategorie (Vokal, CVC-Cluster, Wort, Satz). Abbildung 9.8 auf der nächsten Seite stellt die Leistungen der Gruppen in den unterschiedlichen Kategorien (V, CVC, W, S) genauer dar. Tabelle 9.5 gibt eine Übersicht über die entsprechenden PC-Werte (*percentage correct*) (vgl. auch Abbildung 12.12).

9. Ergebnisse

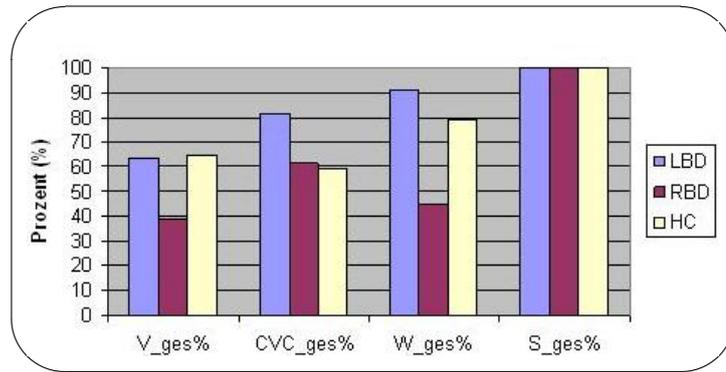


Abbildung 9.8.: Leistungen auf V-, CVC-, W- und S-Ebene (sE + vE)

| | V_ges (%) | CVC_ges (%) | W_ges (%) | S_ges (%) |
|-----|-----------|--------------------|-----------------------|-----------|
| LBD | 63,64 | 81,82 | 90,91 | 100,0 |
| RBD | 38,89 | 61,11 | 44,44 ^{*/**} | 100,0 |
| HC | 64,71 | 58,82 ^t | 79,41 | 100,0 |

Tabelle 9.5.: Leistungen nach Stimuli

Zur Überprüfung etwaiger Innersubjekteffekte wurde ein konservativer F-Test mit angenommener Sphärizität nach Huyhn-Feldt gerechnet. Demnach zeigt sich ein hochsignifikanter Effekt bezüglich des Stimulustyps (V, CVC, W, S [styp]) ($F[3,102] = 17,781$; $p < .001$). Auch die Wechselwirkung *styp* * *Gruppe* hat einen signifikanten Einfluss ($F[6,102] = 2,222$; $p = .047$) (vgl. Abbildung 12.13 auf Seite 204).

Die Überprüfung der Zwischensubjekteffekte zeigt einen hochsignifikanten Effekt ($F[1,34] = 496,542$; $MS_{FEHL} = 1507,571$; $p < .001$, vgl. Abbildung 12.14 auf Seite 204 im Anhang). Diese Zwischensubjekteffekte sollen nun auf den unterschiedlichen Ebenen der Stimulustypen genauer betrachtet werden.

9. Ergebnisse

Vokalebene

Auf V-Ebene zeigen sich im U-Test keine statistischen Unterschiede zwischen den Leistungen der Gruppen LBD, RBD und HC. Die Gruppenleistungen zeigt Tabelle 9.6.

| V | sE (%) | vE (%) |
|----------|--------|--------|
| LBD | 50,0 | 13,64 |
| RBD | 27,78 | 11,11 |
| HC | 44,12 | 20,59 |

Tabelle 9.6.: Leistungen auf Vokalebene

CVC-Ebene

Nach einer Analyse der Leistungen mit Hilfe des U-Tests zeigt sich auf CVC-Ebene zwischen den Gruppen LBD und HC bei der Gesamtleistung $sE + vE$ (81,82 % bzw. 58,82 %; $p = .092$) eine statistische Tendenz. Tabelle 9.7 gibt einen Überblick über die unterschiedlichen sE - bzw. vE -Leistungen (vgl. auch Abbildung 12.15 auf Seite 205 im Anhang).

| CVC | sE (%) | vE (%) |
|------------|--------|--------|
| LBD | 72,73 | 9,09 |
| RBD | 50,0 | 11,11 |
| HC | 52,94 | 5,88 |

Tabelle 9.7.: Leistungen auf CVC-Ebene

9. Ergebnisse

Wortebene

Auf Wortebene unterscheiden sich die Leistungen der Gruppen LBD und RBD in der Variable sE signifikant (86,36 % bzw. 44,44 %; $p = .014$, U-Test nach Mann & Whitney). Ein weiterer signifikanter Unterschied zwischen den beiden Gruppen LBD und RBD zeigt sich bei der Gesamterkennungsleistung ($sE + vE$) (90,91 % bzw. 44,44 %; $p = .006$, vgl. Abbildung 12.16 im Anhang).

Ein Vergleich der Gruppen LBD und HC zeigt hingegen lediglich eine statistische Tendenz der Variable sE (86,36 % bzw. 64,71 %, $p = .051$, vgl. Abbildung 12.17 im Anhang).

Darüber hinaus zeigen sich auf Wortebene zwei statistische Unterschiede zwischen den Gruppen RBD und HC. Zum einen bei der Gesamtleistung $sE + vE$ (44,44 % bzw. 79,41 %, $p = .022$), zum anderen bei der Variable vE (0,00 % bzw. 14,71 %, $p = .076$, vgl. Abbildung 12.18 im Anhang.).

Besonders prominent ist auf Wortebene der (sehr) signifikante Unterschied zwischen der RBD-Gruppe auf der einen und der LBD- bzw. HC-Gruppe auf der anderen Seite (siehe oben). Ein Mittelwertvergleich der RBD-Erkennungsleistung von Stimulus W1 bzw. W2 zeigt keinen signifikanten Unterschied zwischen den beiden Stimuli ($p = .347$).

| W | sE (%) | vE (%) |
|----------|--------------------|------------------|
| LBD | 86,36 | 4,55 |
| RBD | 44,44* | ,00 ^t |
| HC | 64,71 ^t | 14,71 |

Tabelle 9.8.: Leistungen auf Wortebene

9. Ergebnisse

Satzebene

Auf Satzebene zeigen sich keinerlei statistische Unterschiede zwischen den Gruppenleistungen (vgl. Tabelle 9.9).

| S | sE (%) | vE (%) |
|-----|--------|--------|
| LBD | 100,0 | 0 |
| RBD | 94,44 | 5,56 |
| HC | 100,0 | 0 |

Tabelle 9.9.: Leistungen auf Satzebene

9.3.3. Erkennungsleistungen nach Stimulidauer

Nachdem die Erkennungsleistung zunächst aufgrund der linguistischen Kategorie der verwendeten Stimuli überprüft wurde, liegt der Fokus des folgenden Abschnitts auf der Erkennungsleistung in Abhängigkeit von der Stimulidauer. Dabei sollen nun etwaige Längeneffekte anhand der Stimuli CVC1 / CVC2, W1 / W2 bzw. S1 / S2 kontrolliert werden, da alle drei Stimuluspaare zwar der jeweils gleichen linguistischen Kategorie angehören, sich jedoch in ihrer zeitlichen Länge unterscheiden (vgl. Tabelle 9.10). Ein Mittelwertvergleich mit Hilfe des t-Tests für gepaarte Stichproben zeigt keine Korrelationen, jedoch für beide Variablenpaare eine statistische Tendenz (CVC1 — CVC2, $p = .090$ bzw. W1 — W2, $p = .051$) (siehe Abbildung 12.19 und 12.20 auf Seite 206). Ein Zusammenhang zwischen der Stimuluslänge und der Erkennungsleistung konnte demnach nicht nachgewiesen werden.

9. Ergebnisse

| Item | Stimulus | IPA | Dauer (ms) | Mittelwert | SD | p |
|------|----------|------------------------------------------------------------------|------------|--------------------|-------|------|
| 3 | CVC1 | [ˈtan] | 267 | 75,68 ^t | 43,5 | ,090 |
| 4 | CVC2 | [pu:k] | 390 | 56,76 | 50,22 | |
| 5 | W1 | [ˈɔyro:] | 512 | 64,86 | 48,4 | ,051 |
| 6 | W2 | [preˈzi:dent] | 754 | 83,78 ^t | 37,37 | |
| 7 | S1 | [di:frau fyrt das ʊntəˈne:mən ɛgˈfɔlkraɪç] | 2600 | 100 | 0 | |
| 8 | S2 | [ˈy:be drai miˈli:ɔ:n ˈmɛnʃn ˈle:bɔn ɪn deɪgˈhauptʃtat bɛrˈli:n] | 3327 | 100 | 0 | |

Tabelle 9.10.: Übersicht über die verwendeten Stimuli nach Stimulidauer

9.4. Zusammenhänge zwischen Stimmerkennungsleistung und Vortestergebnissen

Mit Hilfe des Tests des Kontingenzkoeffizienten nach Pearson soll untersucht werden, ob sich ein signifikanter Verteilungszusammenhang zwischen unterschiedlichen Messreihen findet. Von besonderer Bedeutung sind dabei Korrelationen zwischen den einzelnen Ergebnissen der Vortests und der Stimmerkennungsleistung im Wahrnehmungsexperiment. Die vollständigen Tabellen finden sich im Anhang ab Seite 207.

9.4.1. Gruppe 1 (LBD)

In der LBD-Gruppe zeigen sich keinerlei statistisch signifikante Korrelationen zwischen den Leistungen in den Vortests und denen des Wahrnehmungsexperimentes. Erwartungsgemäß zeigen sich lediglich statistische Tendenzen zwischen dem Alter der Patienten und den Prozenträngen „R“ (für die Anzahl richtiger Items) und „G“ (für den Gesamtwert, unter Berücksichtigung der Fehlerquote) des AKT (vgl. Abbildung 12.21 auf Seite 207 im Anhang).

9.4.2. Gruppe 2 (RBD)

Die Gruppe der Patienten mit einer rechtsseitigen Läsion zeigt neben den zu erwartenden signifikanten Korrelationen zwischen Lebensalter und den Prozenträngen „T“ (für die Bearbeitungszeit) und „F“ (für die Fehlerzahl) des AKT noch einige weitere statistische Zusammenhänge, die in Tabelle 9.11 aufgeführt werden (vgl. auch Abbildung 12.22 auf Seite 207 im Anhang).

| Variablenpaar | <i>r</i> | <i>p</i> |
|-------------------------------------------|--------------------|----------|
| post-onset time * AKT Prozentrang „R“ | -,756* | ,018 |
| post-onset time * Stimmerkennung (PC) | ,729* | ,026 |
| AKT Prozentrang „R“ * Stimmerkennung (PC) | -,661 [†] | ,053 |
| AKT Prozentrang „G“ * WMS-R (RW) | -,936* | ,039 |

Tabelle 9.11.: Statistisch relevante Korrelationen (RBD)

9. Ergebnisse

9.4.3. Gruppe 3 (HC)

Neben den statistischen Zusammenhängen zwischen dem Probandenalter und dem AKT trat lediglich eine Korrelation zwischen den Variablen *Alter* und *HADS (RW)* auf ($r = .502^*$, $p = .040$) (vgl. auch Abbildung 12.23 auf Seite 208 im Anhang).

9.4.4. Gesamtgruppe BD

Fasst man die Gruppen LBD und RBD zur statistischen Auswertung zu einer gemeinsamen Gruppe zerebral beeinträchtigter Patienten (BD) zusammen, so zeigen sich die folgenden Korrelationen (vgl. Tabelle 9.12 und Abbildung 12.24 auf Seite 209 im Anhang)

| Variablenpaar | <i>r</i> | <i>p</i> |
|-------------------------------------------|----------|----------|
| AKT Prozentrang „T“ * Stimmerkennung (PC) | ,561* | ,010 |
| AKT Prozentrang „F“ * WMS-R (RW) | -,495* | ,026 |
| AKT Prozentrang „G“ * WMS-R (RW) | -,500* | ,025 |

Tabelle 9.12.: Statistisch relevante Korrelationen (BD)

9.4.5. Gesamtgruppe

Überprüft man statistische Zusammenhänge über alle teilnehmenden Patienten und Probanden, so zeigen sich – mit den üblichen Einschränkungen *Alter* * *AKT* – lediglich tendenzielle Zusammenhänge zwischen *Alter* * *WMS-R (RW)*, *AKT Prozentrang* „T“ * *Stimmerkennung (PC)* und *AKT Prozentrang* „F“ * *WMS-R (RW)*, vgl. Abbildung 12.25 auf Seite 210 im Anhang).

9.5. Rating von Schwierigkeitsgrad und subjektiver Sicherheit

Wie in Kapitel 8.4.4 erläutert, wurden die Teilnehmer gebeten, nach jedem Testitem (V1, V2, CVC1, CVC2 etc.) auf einer fünfstufigen Ratingskala (-2 bis +2) einzuschätzen, a) wie sicher sie sich bei der Wahl der gewählten Stimme seien (extrem unsicher – extrem sicher) und b) als wie schwierig sie die jeweilige Aufgabe empfanden (extrem schwer – extrem leicht) (vgl. Abbildungen 13.13 und 13.14).

9.5.1. Gruppe 1 (LBD)

Die folgenden Abbildungen zeigen zunächst die Bewertungen der LBD-Gruppe in Bezug auf den subjektiv empfundenen Schwierigkeitsgrad der jeweiligen Aufgaben (vgl. Abbildung 9.9), das subjektive Sicherheitsgefühl bei der Auswahl der Stimme (vgl. Abbildung 9.10) und der tatsächlich korrekten Erkennungsleistung (vgl. Abbildung 9.11) (vgl. auch Tabellen 12.6 und 12.7 auf Seite 195 im Anhang). Um eine bessere Übersicht zu gewährleisten, wurden die Items in allen Balkendiagrammen in Anlehnung an Tabelle 8.4 auf Seite 116 durchnummeriert.

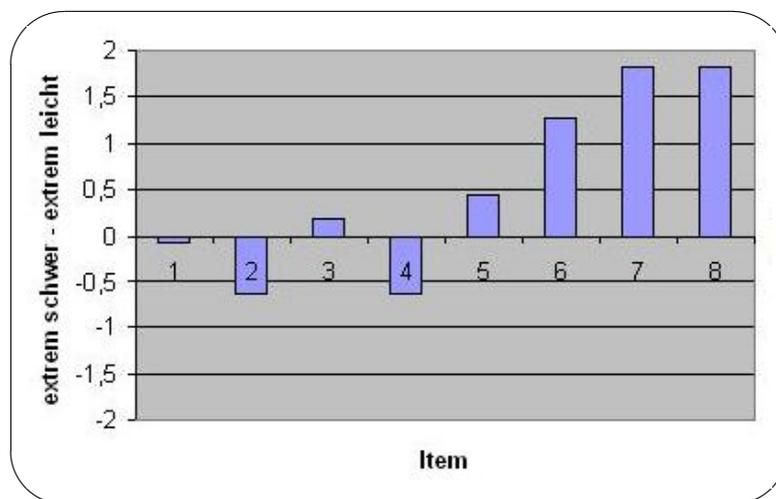


Abbildung 9.9.: Subjektiver Schwierigkeitsgrad (LBD)

1 = V1, 2 = V2, 3 = CVC1, 4 = CVC2, 5 = W1, 6 = W2, 7 = S1, 8 = S2

Um Zusammenhänge zwischen der Bewertung der *subjektiven Schwierigkeit* und dem *subjektiven Sicherheitsgefühl* einer Aufgabe und der tatsächlichen *Erkennungsleistung* identifizieren zu können, wurde der Test des Rangkorrelationskoeffizienten

9. Ergebnisse

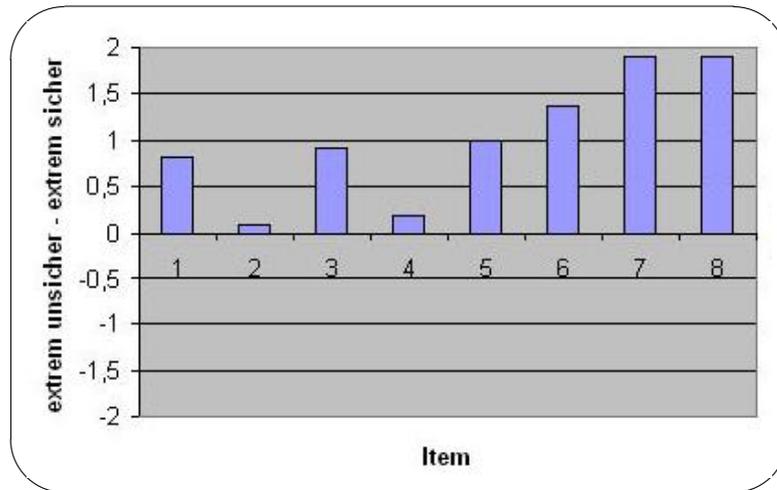


Abbildung 9.10.: Subjektives Sicherheitsgefühl (LBD)

1 = V1, 2 = V2, 3 = CVC1, 4 = CVC2, 5 = W1, 6 = W2, 7 = S1, 8 = S2

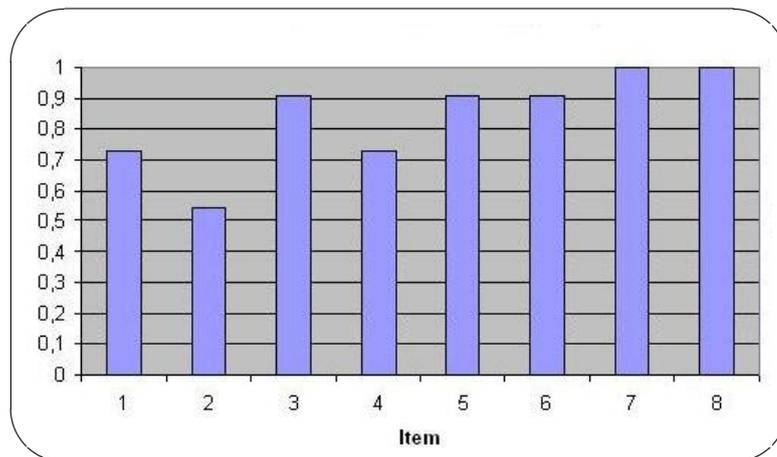


Abbildung 9.11.: Erkennungsleistungen nach Items (LBD)

1 = V1, 2 = V2, 3 = CVC1, 4 = CVC2, 5 = W1, 6 = W2, 7 = S1, 8 = S2

nach Spearman (*Spearman's Rho*) angewendet. Dabei zeigt sich bei Item *V2* ein tendenzieller Zusammenhang zwischen der *Erkennungsleistung* auf der einen und der subjektiv empfundenen *Schwierigkeit* bzw. der subjektiven *Sicherheit* auf der anderen Seite. Ein signifikanter bzw. sehr signifikanter Zusammenhang besteht zwischen der subjektiv empfundenen *Schwierigkeit* und dem subjektiv empfundenen *Sicherheitsgefühl* bei den Stimuli *V1* und *V2*. Ein tendenzieller Zusammenhang zwischen der *Erkennungsleistung* und der *subjektiven Sicherheit* konnte für Item *W2* nachgewiesen werden. Einen signifikanten Zusammenhang gibt es auf W-Ebene zwischen der empfundenen *Schwierigkeit* und dem empfundenen *Sicherheitsgefühl* bei Stimulus *W2* und schließlich auf Satzebene (*S1* und *S2*) ebenfalls zwischen dem subjektiven

9. Ergebnisse

Schwierigkeits- und dem entsprechenden *Sicherheitsgefühl*. Darüber hinaus scheinen die LBD-Patienten die Items *S1* und *S2* als gleich leicht bzw. schwer empfunden zu haben, da sich eine hochsignifikante Korrelation zwischen den Variablen *Schwierigkeit X1 * Schwierigkeit X2* bzw. *Sicherheit X1 * Sicherheit X2* nachweisen lässt (bei beiden $r = 1.00^{***}$, $p < .001$). Sämtliche Korrelationen zeigen auch die Abbildungen 12.26 bis 12.29 ab Seite 211 im Anhang.

| Item | Variablenpaar | <i>r</i> | <i>p</i> |
|------|------------------------------------|-------------------|----------|
| V1 | Schwierigkeit * Sicherheit | ,641* | ,034 |
| V2 | Schwierigkeit * Sicherheit | ,772** | ,005 |
| V2 | Erkennungsleistung * Schwierigkeit | ,577 ^t | ,063 |
| | Erkennungsleistung * Sicherheit | ,581 ^t | ,061 |
| W2 | Erkennungsleistung * Sicherheit | ,553 ^t | ,078 |
| | Schwierigkeit * Sicherheit | ,716* | ,013 |
| S1 | Schwierigkeit * Sicherheit | ,671* | ,024 |
| S2 | | ,671* | ,024 |

Tabelle 9.13.: Ausgewählte Korrelationen der Ratingskalen (LBD)

9.5.2. Gruppe 2 (RBD)

Die folgenden Abbildungen (9.12, 9.13 und 9.14) geben zunächst die Bewertungen der RBD-Gruppe in Bezug auf den subjektiv empfundenen *Schwierigkeitsgrad* der jeweiligen Aufgaben, das subjektive *Sicherheitsgefühl* bei der Auswahl der Stimme und der tatsächlich korrekten *Erkennungsleistung* wieder (vgl. auch Tabellen 12.8 und 12.9 auf Seite 198 im Anhang).

Ähnlich wie auch bei der LBD-Gruppe zeigen sich einige (sehr) signifikante Korrelationen zwischen dem Rating der Teilnehmer und ihrer tatsächlichen *Erkennungsleistung*. So besteht bei einem Korrelationskoeffizienten von $r = .880^{**}$ ein sehr signifikanter Zusammenhang zwischen *Erkennungsleistung* und *Subjektive Schwierigkeit* bei Item *V1* ($p = .002$) sowie ein signifikanter Zusammenhang ($r = .755^*$, $p = .019$) mit *Subjektiver Sicherheit*. Darüber hinaus zeigt sich ein signifikanter Zusammenhang zwischen *Erkennungsleistung* und *Subjektiver Sicherheit* bei Item *W1* ($r = .732^*$, $p = .025$; vgl. Tabelle 9.15). Wie bereits bei der LBD-Gruppe zeigt auch die RBD-Gruppe mehrere signifikante Korrelationen zwischen den Variablen

9. Ergebnisse

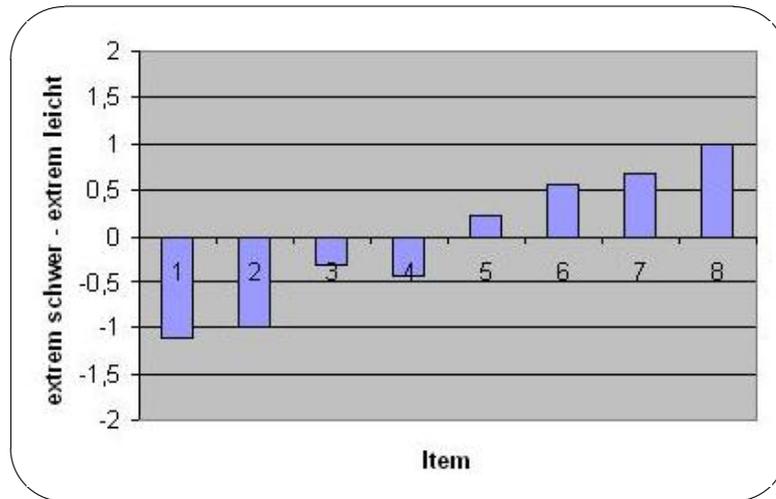


Abbildung 9.12.: Subjektiver Schwierigkeitsgrad (RBD)

1 = V1, 2 = V2, 3 = CVC1, 4 = CVC2, 5 = W1, 6 = W2, 7 = S1, 8 = S2

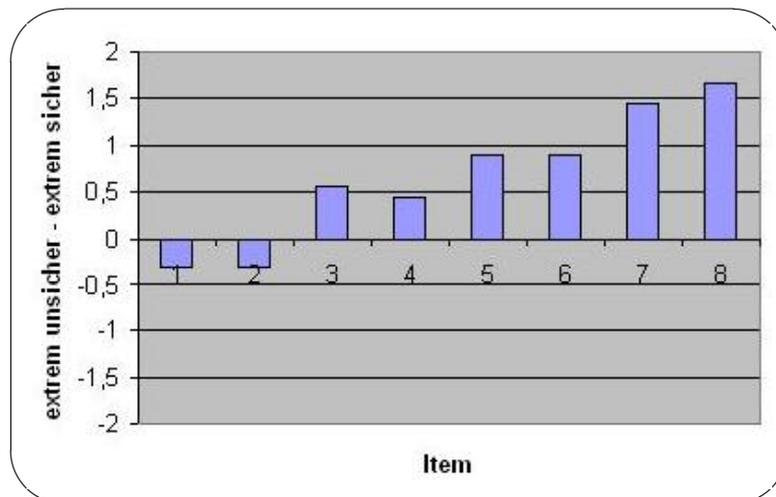


Abbildung 9.13.: Subjektives Sicherheitsgefühl (RBD)

1 = V1, 2 = V2, 3 = CVC1, 4 = CVC2, 5 = W1, 6 = W2, 7 = S1, 8 = S2

*Schwierigkeit * Sicherheit* bei den Items V1, V2, W2, S1 und S2. Die Zusammenhänge u.a. zwischen den Variablen *Schwierigkeit X1 * Schwierigkeit X2* zeigen auch bei dieser Gruppe, dass die einzelnen Itemgruppen für sich (V1-V2, CVC1-CVC2, W1-W2, S1-S2) als gleich schwierig empfunden wurden. Eine vollständige Übersicht über aller Kovarianzen zeigen die Abbildungen 12.30 bis 12.33 ab Seite 213 im Anhang.

9. Ergebnisse

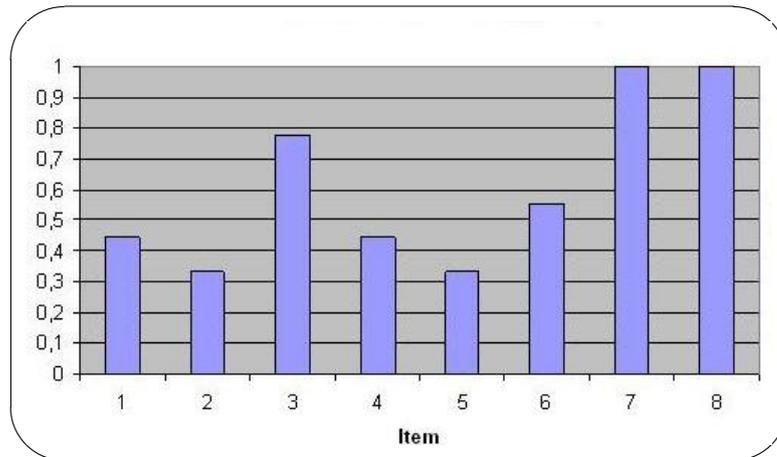


Abbildung 9.14.: Erkennungsleistungen nach Items (RBD)

1 = V1, 2 = V2, 3 = CVC1, 4 = CVC2, 5 = W1, 6 = W2, 7 = S1, 8 = S2

| Item | Variablenpaar | r | p |
|------|------------------------------------|-------------------|------|
| V1 | Erkennungsleistung * Schwierigkeit | ,880** | ,002 |
| | Erkennungsleistung * Sicherheit | ,755* | ,019 |
| | Schwierigkeit * Sicherheit | ,667* | ,05 |
| V2 | Schwierigkeit * Sicherheit | ,618 ^t | ,076 |
| W1 | Erkennungsleistung * Sicherheit | ,732* | ,025 |
| W2 | Schwierigkeit * Sicherheit | ,690* | ,040 |
| S1 | Schwierigkeit * Sicherheit | ,704* | ,034 |
| S2 | | ,686* | ,041 |

Tabelle 9.14.: Ausgewählte Kovarianzen der Ratingskalen (RBD)

9.5.3. Gruppe 3 (HC)

Die folgenden Abbildungen geben die Ergebnisse der Ratings und die tatsächliche *Erkennungsleistung* der zerebral gesunden Kontrollprobanden wieder. Die entsprechenden Messwerte dazu zeigen Tabellen 12.10 und 12.11 im Anhang auf Seite 197.

Eine Überprüfung etwaiger Zusammenhänge mit Hilfe des *Spearman's Rho* ergab lediglich eine sehr signifikante Korrelation zwischen den Variablen *Erkennungsleistung* und *Subjektive Sicherheit* bei Item CVC1 ($r = .611^{**}$, $p = .009$, vgl. Tabelle 9.15 auf Seite 144). Eine statistische Tendenz zeigte sich bei Stimulus W2 zwischen den Variablen *Erkennungsleistung* und *Schwierigkeit* bzw. *Sicherheit*. Wie bereits bei den anderen Gruppe zuvor berichtet, empfanden auch die HC-Probanden einige Itempaare als gleich schwierig. Dies gilt für die Paare V1-V2, W1-W2 und

9. Ergebnisse

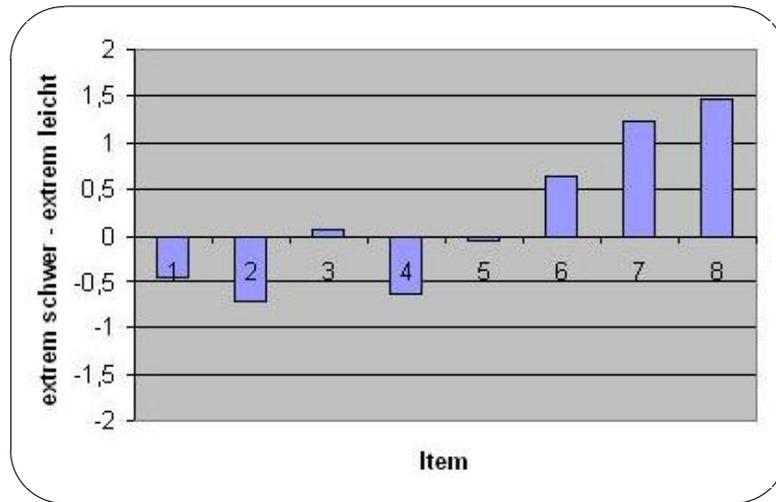


Abbildung 9.15.: Subjektiver Schwierigkeitsgrad (HC)

1 = V1, 2 = V2, 3 = CVC1, 4 = CVC2, 5 = W1, 6 = W2, 7 = S1, 8 = S2

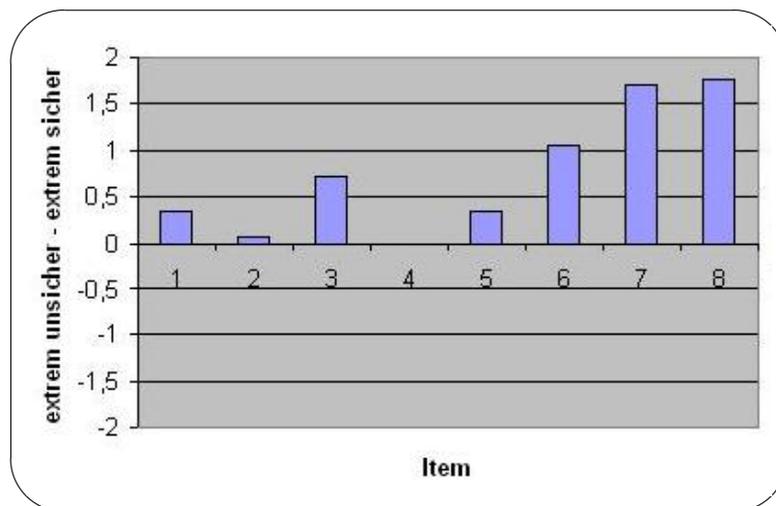


Abbildung 9.16.: Subjektives Sicherheitsgefühl (HC)

1 = V1, 2 = V2, 3 = CVC1, 4 = CVC2, 5 = W1, 6 = W2, 7 = S1, 8 = S2

S1–S2. Eine Übersicht über sämtliche Korrelationen geben die Abbildungen 12.34 bis 12.37 ab Seite 215 im Anhang.

9.5.4. Gesamtgruppe

Wie bereits bei den oben beschriebenen Gruppen sollen auch hier zunächst die Abbildungen 9.18 bis 9.20 die Bewertungen der Gesamtgruppe (LBD + RBD + HC) in Bezug auf den subjektiv empfundenen *Schwierigkeitsgrad* der jeweiligen Aufgaben, das subjektive *Sicherheitsgefühl* bei der Auswahl der Stimme und der tatsächlich kor-

9. Ergebnisse

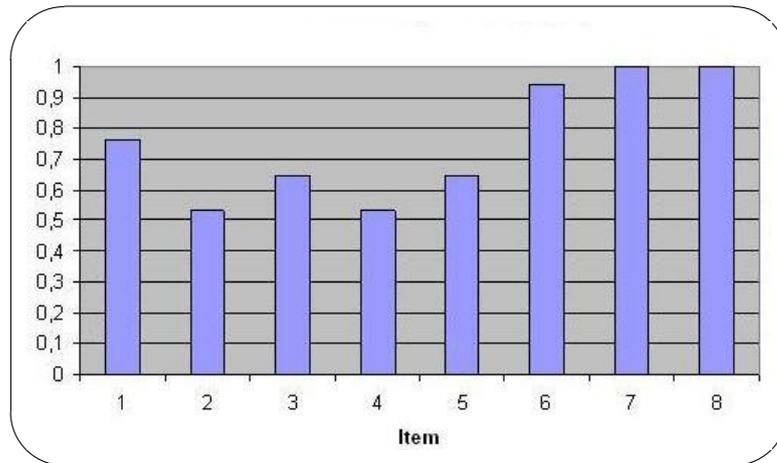


Abbildung 9.17.: Erkennungsleistungen nach Items (HC)

1 = V1, 2 = V2, 3 = CVC1, 4 = CVC2, 5 = W1, 6 = W2, 7 = S1, 8 = S2

| Item | Variablenpaar | r | p |
|-------|------------------------------------|-------------------|------|
| CVC 1 | Erkennungsleistung * Sicherheit | ,611** | ,009 |
| W2 | Erkennungsleistung * Schwierigkeit | ,450 ^t | ,70 |
| | Erkennungsleistung * Sicherheit | ,414 ^t | ,098 |

Tabelle 9.15.: Ausgewählte Kovarianzen der Ratingskalen (HC)

rekten *Erkennungsleistung* graphisch darstellen (vgl. auch Tabellen 12.12 und 12.13 auf Seite 198 im Anhang).

Auf die Gesamtpopulation gerechnet zeigt der *Spearman's Rho* mehrere signifikante Korrelationen. So lässt sich ein (sehr) signifikanter Zusammenhang zwischen den Variablen *Erkennungsleistung* und *Subjektive Schwierigkeit* bei Item *V1* ebenso nachweisen ($r = .509^{**}$, $p = .001$), wie zwischen *Erkennungsleistung* und *Sicherheit* bei Item *V2* ($r = .383^*$, $p = .019$). Auch bei *CVC1* zeigt sich ein Zusammenhang zwischen *Erkennungsleistung* und *Subjektive Sicherheit* ($r = .484^{**}$, $p = .002$) sowie bei *CVC2* zwischen *Erkennungsleistung* und *Schwierigkeit* ($r = .334^*$, $p = .044$). Auf Wortebene (*W2*) zeigen sich sehr signifikante Zusammenhänge zwischen *Erkennungsleistung* und *Subjektive Schwierigkeit* ($r = .476^{**}$, $p = .003$) bzw. *Subjektive Sicherheit* ($r = .486^{**}$, $p = .002$). Auch über die Gesamtpopulation zeigt sich, dass sämtliche einzelne Itempaare als gleich schwierig empfunden wurden (*V1-V2*, *CV1-CVC2*, *W1-W2* und *S1-S2*). Eine Übersicht über die hier genannten Korrelationen gibt Tabelle 9.16 auf der nächsten Seite. Die restlichen Zusammenhänge zeigen die Abbildungen 12.38 bis 12.41 ab Seite 217 im Anhang.

9. Ergebnisse

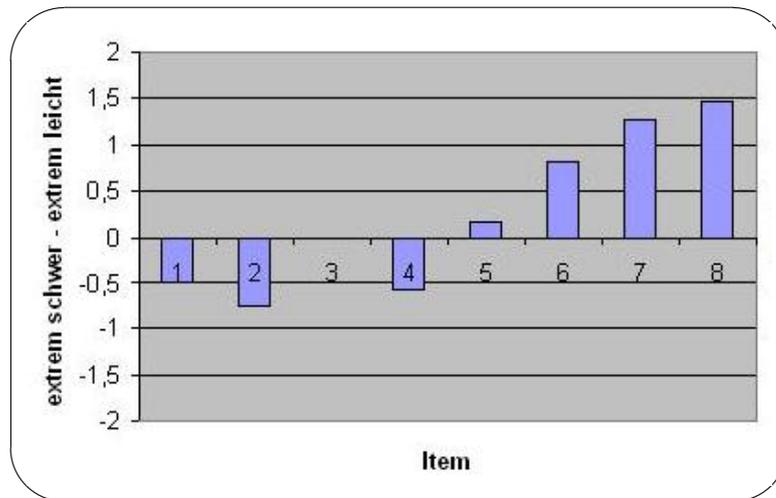


Abbildung 9.18.: Subjektiver Schwierigkeitsgrad (Gesamt)

1 = V1, 2 = V2, 3 = CVC1, 4 = CVC2, 5 = W1, 6 = W2, 7 = S1, 8 = S2

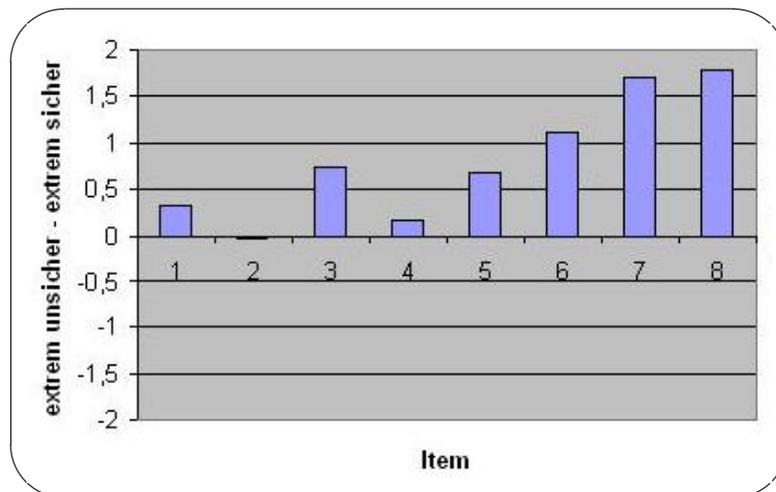


Abbildung 9.19.: Subjektives Sicherheitsgefühl (Gesamt)

1 = V1, 2 = V2, 3 = CVC1, 4 = CVC2, 5 = W1, 6 = W2, 7 = S1, 8 = S2

| Item | Variablenpaar | <i>r</i> | <i>p</i> |
|------|------------------------------------|----------|----------|
| V1 | Erkennungsleistung * Schwierigkeit | ,880** | ,002 |
| V2 | Erkennungsleistung * Sicherheit | ,383* | ,019 |
| CVC1 | Erkennungsleistung * Sicherheit | ,484** | ,002 |
| CVC2 | Erkennungsleistung * Schwierigkeit | ,334* | ,044 |
| W2 | Erkennungsleistung * Schwierigkeit | ,476** | ,003 |
| | Erkennungsleistung * Sicherheit | ,486** | ,002 |

Tabelle 9.16.: Ausgewählte Kovarianzen der Ratingskalen (Gesamt)

9. Ergebnisse

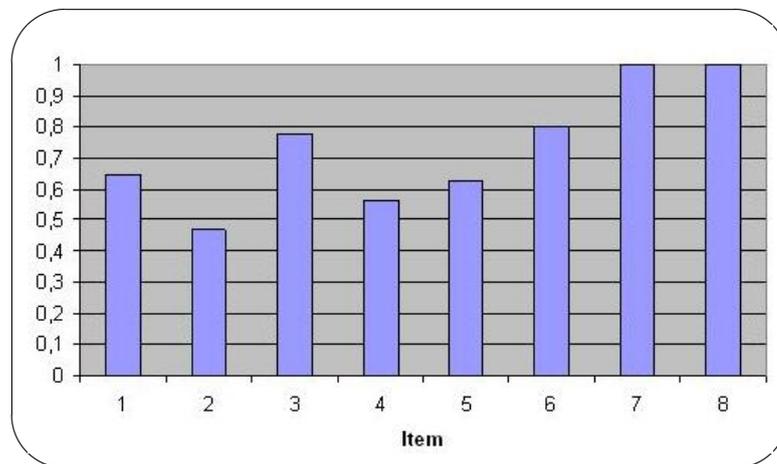


Abbildung 9.20.: Erkennungsleistungen nach Items (Gesamt)
1 = V1, 2 = V2, 3 = CVC1, 4 = CVC2, 5 = W1, 6 = W2, 7 = S1, 8 = S2

10. Zusammenfassung und Diskussion

10. Zusammenfassung und Diskussion

Im Mittelpunkt der vorliegenden Untersuchung steht die Fragestellung, inwiefern aphasische bzw. nicht-aphasische Patienten in der Lage sind, die Stimmen einer sehr vertrauten Person aus einer Auswahl mehrerer Stimmuster zu erkennen. Zu diesem Zweck wurden drei Probandengruppen (LBD, RBD und HC) einer quasiexperimentellen Wahrnehmungsuntersuchung unterzogen, in deren Verlauf sie im Rahmen von acht unterschiedlichen Aufgabenstellungen die Stimme des Ehe- / Lebenspartners unter sechs Stimmustern erkennen sollten.

Im Verlaufe dieses Kapitels werden nun die wichtigsten Untersuchungsergebnisse zusammengefasst und diskutiert. Dabei soll sich vor allem an den in Kapitel 8.2 aufgestellten Hypothesen orientiert werden.

Wie in Kapitel 9.3 beschrieben, konnte die Hypothese bestätigt werden, dass Patienten mit einer rechtshemisphärischen Läsion im Leistungsbereich der auditiven Stimmerkennung den durchschnittlichen Leistungen der LBD-Patienten bzw. HC-Probanden unterlegen sind (vgl. Tabelle 9.4 und Abbildung 9.7 auf Seite 128).

Die Korrelationen zwischen Defiziten bei der Wiedererkennung und einer Läsion innerhalb der rechten Hemisphäre, insbesondere des rechten Parietallappens bzw. Temporallappens, bestätigen die bisherigen Theorien zur Funktion der rechten Hemisphäre, die ihr die Verarbeitung nonlinguistischer, prosodischer Informationen zuschreiben (vgl. u.a. Pell 2006, Homae et al. 2006, Lattner, Meyer & Friederici 2005, Wildgruber et al. 2002, 2005, Mitchell et al. 2003, Scott et al. 1997, Kapur et al. 1992, 1994, Zatorre et al. 1992, Ellis, Young & Critchley 1989). Die Leistungsunterschiede zwischen den RBD-Patienten und v.a. den LBD-Patienten unterstützen darüber hinaus die bereits ausführlich in Kapitel 7.3.2 ab Seite 88 vorgestellten Untersuchungsergebnisse, die die Verarbeitung von Stimmustern eher der rechten Großhirnhemisphäre zuschreiben (vgl. u.a. Belin, Fecteau & Bédard 2004, von Kriegstein & Giraud 2004, Belin & Zatorre 2003 etc.).

Darüber hinaus finden diese Ergebnisse auch Unterstützung in zahlreichen weiteren Studien, die die Verarbeitung anderer komplexer nonverbaler akustischer Muster und ihrer frequenzabhängigen Informationen (Tonhöhen (*pitch*), Tonhöhenkonturen etc.) zum Untersuchungsgegenstand haben. Dazu zählt z.B. die Verarbeitung von Melodien oder Geräuschen. So zeigen zahlreiche Studien zur Musikererkennung übereinstimmend, dass v.a. bei der Verarbeitung frequenzabhängiger Melodien die rechte Hemisphäre in wesentlichem Umfang involviert ist (siehe dazu Vignolo 2003). Bei den beteiligten Strukturen fallen vor allem Aktivierungen im rechtsseitigen *Gyrus*

10. Zusammenfassung und Diskussion

temporalis superior (vgl. u.a. Liégeois-Chauval et al. 1998:1859f, Peretz 1990:1193f), im anterolateralen Anteil der Heschl'schen Querwindung (vgl. u.a. Warrier & Zatorre 2004:1619f, Tramo, Shah & Braidà 2002:128f, Johnsrude, Penhune & Zatorre 2000:160) sowie – das zeigen bildgebende Verfahren – in dem an die Heschl'sche Querwindung angrenzenden posterioren Anteil der sekundären Hörrinde auf (vgl. u.a. Hart, Palmer & Hall 2003, Hall et al. 2002). Entsprechend zeigen Patienten mit Läsionen der genannten Regionen Defizite im Bereich der Melodieerkennung und der -diskrimination (vgl. u.a. Ayotte et al. 2000:1931f).

Da auch das Erkennen und Verstehen affektiver und linguistischer Prosodie maßgeblich auf der Verarbeitung und Interpretation von Frequenzverläufen beruht, weisen Patienten mit den eben beschriebenen Läsionen auch Auffälligkeiten und Defizite in diesem Bereich der Verarbeitung akustischer Parameter auf (vgl. dazu u.a. Pell 2006, 1999, 1998, Pell et al. 1997, George et al. 1996, Van Lancker & Sidtis 1992). Zusammenfassend lässt sich also festhalten, dass spezifische Verletzungen in der rechten Großhirnhemisphäre häufig mit Defiziten bei der auditiven Verarbeitung (Erkennung, Diskriminierung) akustischer Muster korrelieren. Dies kann die Verarbeitung von Stimmen genauso betreffen, wie die Verarbeitung von Melodien und Prosodie. Dass Amusie und Aprosodie auch parallel einher gehen können, zeigen die Fallbeschreibungen von Bautista & Ciampetti (2003) oder Confavreux et al. (1998) (siehe dazu auch Peretz 2001:535).

Die in Kapitel 8.2 aufgestellte Hypothese, die die Defizite der RBD-Patienten postulierte, konnte anhand der vorliegenden Ergebnisse somit bestätigt werden.

Trotz der zahlreichen Hinweise in der Literatur muss an dieser Stelle jedoch angemerkt werden, dass die Leistungsunterschiede zwischen der RBD- und der HC-Gruppe bei weitem nicht so prominent ausgefallen sind, wie auf Grund der in der Literatur beschriebenen Untersuchungen vermutet werden konnte. Da es keinen Anlass gibt, die Ergebnisse der Kontrollprobanden als Durchschnittsleistung unbeeinträchtigter Normalpersonen anzuzweifeln, soll an dieser Stelle diskutiert werden, weshalb die Leistungen der HC- und der RBD-Gruppe statistisch beinahe gleich sind und sich lediglich tendenziell voneinander unterscheiden ($p = .093$).

Ein Grund liegt sicherlich darin, dass aufgrund der ungünstigen Rahmenbedingungen¹ davon ausgegangen werden muss, dass bei einzelnen RBD-Patienten die für die Stimmerkennung kritischen Areale (v.a. der *Sulcus temporalis superior* etc. [vgl. Kapitel 7.3.2]) nicht oder nur zu einem geringen Teil von der Läsion betroffen waren.

¹Hiermit ist die mangelnde Kooperationsbereitschaft der Akutkliniken gemeint, die das Sichten der CT-Scans nicht ermöglichten (vgl. u.a. Kapitel 8.3.1).

10. Zusammenfassung und Diskussion

Geht man davon aus, dass die RBD-Patienten, deren Erkennungsrate über der durchschnittlichen Leistung der HC-Probanden liegt (75,74 %)², keine bzw. lediglich eine geringe Läsion der kritischen Areale aufweisen, so fallen die Leistungsunterschiede zwischen den genannten Gruppen erwartungsgemäß größer aus. Die nach diesen Kriterien ausgewählte RBD-Gruppe (N = 7) weist dann nur noch eine durchschnittliche Erkennungsrate von 51,79 % auf.

Die einfaktorielle ANOVA zeigt selbstverständlich auch in diesem Fall einen sehr signifikanten Unterschied zwischen den einzelnen Gruppen ($F[2,32] = 7,393, p = .002$). Im post-hoc-Test zeigt sich ein (sehr) signifikanter Unterschied zwischen den Leistungen der Gruppen RBD und LBD ($p = .001$) sowie zwischen den Gruppen RBD und HC ($p = .011$). Dieses Vorgehen ist jedoch rein spekulativ und darf keinesfalls verallgemeinert werden.

Nicht nur bei den beiden o.g. RBD-Patienten 12 und 15, deren Leistungen über den durchschnittlichen Leistungen der HC-Probanden lagen, sondern auch bei allen anderen Patienten und Probanden muss berücksichtigt werden, dass die Leistungen im Bereich der Stimmerkennung deutlich von der Qualität der Ziel- bzw. Distraktorstimmen abhängig sind. So ist es nicht auszuschließen, dass sich z.B. die Zielstimmen der Probanden 12 und 15 in ihrer Qualität objektiv wesentlich deutlicher von den Distraktorstimmen unterscheiden als die Zielstimmen anderer Patienten / Probanden. Dieses Argument wurde bei den veröffentlichten Studien zur Phonagnosie bisher nicht berücksichtigt.

Darüber hinaus ist es denkbar, dass die RBD-Patienten ihre Stimmauswahl nicht aufgrund eines unbeeinträchtigten auditiven Analyseprozesses getroffen haben, sondern über ein Ausschlussverfahren.

Der Erläuterung dieses vermuteten Ausschlussverfahrens soll das bereits in Kapitel 7.3.2 vorgestellte Arbeitsmodell zur Personenerkennung von Neuner & Schweinberger (2000:345f) in Anlehnung an Bruce & Young (1986) zugrunde gelegt werden (hier: *voice recognition*; vgl. das modifizierte Modell in Abbildung 10.1).

Dieses Modell sieht vor, dass nach der Stufe des *Encoding of Sound* die Verarbeitungsstufe des *Structural Encoding* folgt, auf der jene Informationen aus dem Stimmsignal extrahiert werden, die als stabil und invariant gelten und nicht von situativen Faktoren beeinflusst werden. Diese individuell spezifischen Muster werden im Anschluss mit jenen verglichen, die als *Voice Recognition Units (VRU)* im episodischen Gedächtnis gespeichert sind. War der Vergleich erfolgreich, so werden die

²Dies betraf die bereits weiter oben vorgestellten Probanden 12 (87,5 %) und 15 (100 %).

10. Zusammenfassung und Diskussion

Person Identity Nodes (PIN) aktiviert. Auf dieser Verarbeitungsstufe ist der Patient in der Lage, ein Urteil über die vermeintliche Bekanntheit der Person zu fällen (vgl. Burton, Bruce & Johnston 1990:361). Wie in Kapitel 7.3.2 bereits angedeutet, können die Stimuli bei auditiven Erkennungsaufgaben nicht parallel, sondern ausschließlich nacheinander dargeboten werden. Aus diesem Grund kann davon ausgegangen werden, dass der Patient nach jedem gehörten Stimulus einen Abgleich mit den gespeicherten *VRUs* vornimmt und eine bekannt / unbekannt-Entscheidung trifft. Es kann also abschließend nicht sicher beantwortet werden,

- ob der Patient tatsächlich auf weitere semantische Informationen („Dieses Stimmuster gehört zu meiner Frau.“) zugreifen konnte (*Retrieval of Semantic Information*) oder
- ob er lediglich die fünf Distraktorstimmen als „unbekannt“ beurteilt hat und seine Wahl auf die einzige „bekannte“ Stimme gefallen ist, ohne dass er möglicherweise definitiv wusste, dass dies die Stimme seiner Ehefrau oder die einer anderen vertrauten Person gewesen ist.

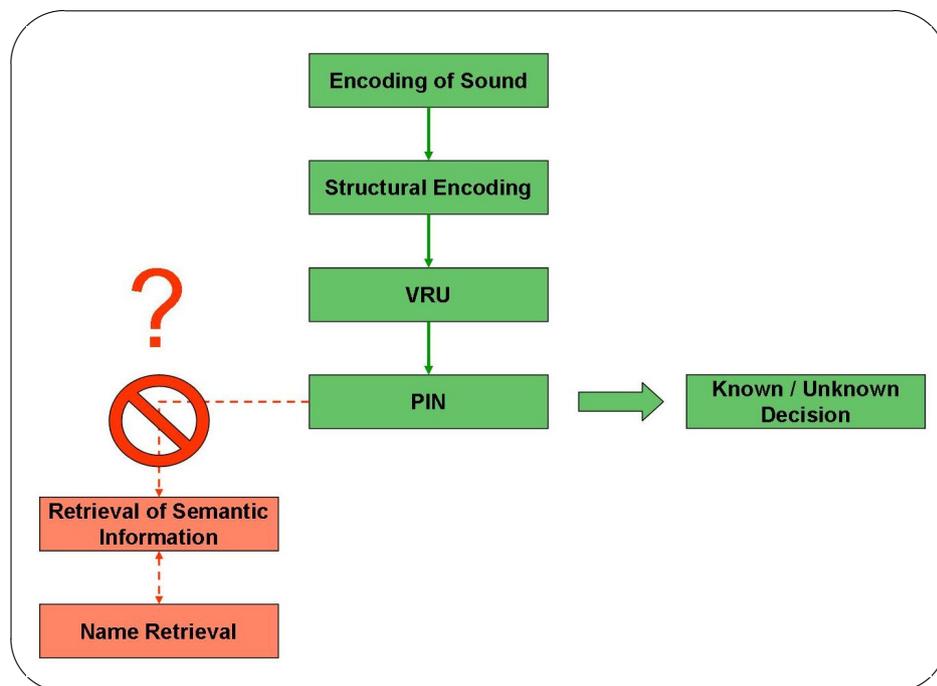


Abbildung 10.1.: Stimmerkennungsmodell in Anlehnung an Bruce & Young (1986).

10. Zusammenfassung und Diskussion

In letztgenanntem Fall wäre es also möglich, dass die RBD-Patienten zwar über Defizite im Bereich der Stimmerkennung verfügen, diese jedoch kompensieren konnten, da sie ihre korrekte Stimmauswahl über den Umweg eines Ausschlussverfahren erreichen konnten.

Auch wenn der Leistungsunterschied zwischen den Gruppen LBD und HC statistisch nicht nachweisbar ist, so überrascht dennoch die unerwartete *numerische* Überlegenheit der aphasischen LBD-Patienten. Während die gesunden Kontrollprobanden lediglich 6,1 der 8 Items korrekt erkannten, waren dies in der LBD-Gruppe durchschnittlich 6,7 Items³. 36,4 % (N = 4) der LBD-Patienten erreichten dabei eine Wiedererkennungsrate von 100 %, was lediglich 29,4 % (N = 5) der HC-Probanden und nur 11,11 % (N = 1) der RBD-Patienten gelang.

Da sich sämtliche LBD-Patienten im Unterschied zu den HC-Probanden zum Testzeitpunkt in intensiver sprachtherapeutischer Behandlung im Rahmen einer AHB befanden bzw. bis kurz zuvor befunden haben, ist es denkbar, dass sie dadurch eine – möglicherweise auch nur vorübergehende – besondere 'Affinität' bzw. Aufmerksamkeit für Sprachstimuli entwickelt haben, die als eine Art 'Priming' verstanden werden könnte. Ob sich ein solches Phänomen während bzw. im Anschluss an eine (intensive) sprachtherapeutischen Behandlung beobachten lässt, wurde in der Literatur bisher noch nicht berichtet.

Als eine weitere Erklärung für die (numerische!) Überlegenheit der LBD-Gruppe wird angenommen, dass insbesondere aphasische Patienten prosodische Informationen des Sprachsignals zur Kompensation der meist im Bereich des auditiven Sprachverständnisses vorhandenen Defizite nutzen. Aphasische LBD-Patienten richten bei der Rezeption gesprochener Sprache ihre Aufmerksamkeit also eher auf die prosodischen Informationen des Sprachsignals, um sich mit deren Hilfe den Sinn des Gesagten leichter erschließen zu können. Sie sind somit aufmerksamer und sensibler für diese Art von Informationen als beispielsweise gesunde Kontrollpersonen, die ihren Fokus unwillkürlich eher auf den semantischen Inhalt des Gesprochenen richten.

Erste Überlegungen dazu, dass Prosodie das auditive Sprachverständnis aphasischer Patienten positiv beeinflussen könnte, stammen bereits aus den sechziger und siebziger Jahren von DeRenzi & Vignolo (1962), Goodglass et al. (1967) sowie von Boller & Green (1972). Bestätigung fand diese Vermutung u.a. in den Untersuchungen von Kimelman (1991, 1999, siehe auch Raithel 2005). In seiner Untersuchung

³Zum Vergleich: in der Gruppe der RBD-Patienten liegt der Wert bei 4,9 Items.

10. Zusammenfassung und Diskussion

von 1999 konfrontierte Kimelman 36 Probanden (26 Aphasiker, 10 Kontrollprobanden) mit insgesamt acht unterschiedlichen Texten, von denen vier *Low-Linguistic-Complexity (LLC)*⁴ und vier *High-Linguistic-Complexity (HLC)*⁵ aufwies. Jeder Text bestand aus acht Sätzen, in denen jeweils ein Zielwort enthalten war. Alle Patienten und Probanden wurden im Verlaufe von zwei Sitzungen mit beiden Texten (LLC bzw. HLC) konfrontiert, während die Zielwörter einmal besonders betont, ein andermal neutral artikuliert wurden. Im Anschluss wurden den Patienten und Probanden Fragen zu den Texten gestellt, deren korrekte Antwort in dem (un-) betonten Zielwort bestand. Die Ergebnisse zeigen, dass alle aphasischen Patienten von der Betonung des Zielwortes profitieren konnten (vgl. Kimelman 1999:217f).

Die Ergebnisse der eben geschilderten Untersuchung und die zahlreicher weiterer bestätigen, dass v.a. aphasische Patienten mit Hilfe der prosodischen Informationen des Sprechschalls ihr auditives Sprachverständnis unterstützen können. Da das Analysieren und Verarbeiten von Frequenzverläufen nicht nur für das Verständnis von Prosodie, sondern auch für das Erkennen von Stimmen von großer Bedeutung ist, kann davon ausgegangen werden, dass die Patienten – sozusagen im „Nebeneffekt“ – auch bei der Stimmerkennung von dieser Strategie profitieren und aus diesem Grund die numerisch beste Leistung der drei Untersuchungsgruppen zeigen.

Belin, Fecteau & Bédard bezeichneten in einem Übersichtsartikel aus dem Jahre 2004 die Stimme einer Person als deren „auditorisches Gesicht“ („auditory face“). Damit spielen sie darauf an, dass das Wiedererkennen von Stimmen nach einem ganz ähnlichen Modell wie das Wiedererkennen von Gesichtern organisiert sein könnte. Denn auch bei der Gesichtererkennung reicht es nicht aus, das Gesehene als menschliches Gesicht zu erkennen, sondern der Betrachter muss es von anderen Exemplaren dieser Gattung unterscheiden können. Natürlich wäre es zunächst grundsätzlich denkbar, von jedem uns vertrauten Gesicht möglichst viele Ansichten zu speichern, um dann während des Wiedererkennungsprozesses die Ansicht zu suchen, die der aktuellen Wahrnehmung am ähnlichsten ist. Dies hätte jedoch den Nachteil, dass wir a) unökonomisch viele Informationen speichern, und b) die abgespeicherten Daten ständig aktualisieren müssten, da sich menschliche Gesichter, z.B. durch Mimik, unterschiedliche Lichtverhältnisse, sich ändernde Paraphernalien (Frisur, Brille etc.) und auch Alterungsprozesse, nahezu im ständigen Wandel befinden. Die Alternative dazu wäre, nur die charakteristischen Merkmale eines Gesichtes zu speichern, die für

⁴Entspricht dem Lese-Level der fünften bis sechsten Klasse.

⁵Entspricht dem Lese-Level der achten bis neunten Klasse.

10. Zusammenfassung und Diskussion

eine zuverlässige Unterscheidung ausreichen sollten und die auch bei wechselnden Sichtwinkeln oder Lichtverhältnissen konstant bleiben (vgl. Goldenberg ⁵2002a:308f, ³2002b:155).

Während es zur auditiven Wiedererkennung akustischer Muster *vergleichsweise* wenige Untersuchungen gibt, finden sich in der Literatur wesentlich mehr Hinweise darauf, dass die visuelle Wiedererkennung optischer Muster bei Patienten mit einer rechtshemisphärischen Läsion eingeschränkt zu sein scheint.

Im Rahmen des im Verlauf der Vortestungen durchgeführten AKT (Gatterer 1990) muss, wie in Kapitel 9.3 bereits beschrieben, ein vorgegebenes Muster mit anderen Mustern visuell verglichen werden (siehe Abbildung 10.2). Aufgrund der Hinweise in der Literatur wurde die Hypothese formuliert, dass die Patienten mit einer rechtshemisphärischen Läsion im Bereich der visuellen Mustererkennung den LBD-Patienten und den HC-Probanden unterlegen sein sollten.

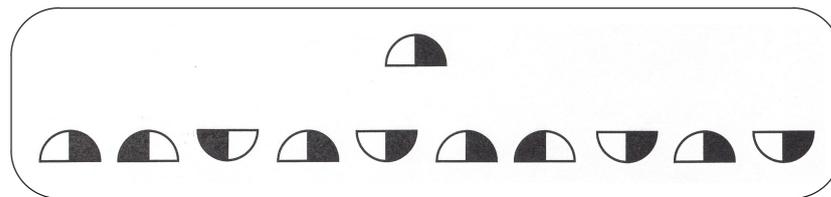


Abbildung 10.2.: Ausschnitt aus einem Testbogen des AKT
(aus: Gatterer 1990)

Um die oben erläuterte Hypothese bestätigen zu können, wurde zunächst ermittelt, ob sich ein statistischer Unterschied zwischen den visuellen Wiedererkennungsleistungen der drei Gruppen nachweisen ließe. Eine Analyse der Messdaten mit Hilfe der einfaktoriellen ANOVA zeigte zunächst einen signifikanten Unterschied zwischen den beteiligten Gruppen ($F[2,34] = 3,904$, $p = .030$). Bei der anschließenden Varianzanalyse konnte nachgewiesen werden, dass der AKT-Prozentrang „R“ der RBD-Gruppe mit 73,54 % signifikant unterhalb der Prozentränge der LBD- (90,53 %, $p = .033$) bzw. HC-Gruppe (88,83 %) liegt ($p = .026$).

Somit konnte mit Hilfe der bereits im vorigen Kapitel erläuterten Ergebnisse die Hypothese bestätigt werden, dass Patienten mit einer rechtshemisphärischen Läsion Defizite im Bereich der visuellen Mustererkennung aufzeigen.

Aufgrund der statistisch nur tendenziellen Zusammenhänge zwischen AKT-Prozentrang „R“ und der Stimmerkennungsleistung, sowie der relativ großen Streuung der AKT-Prozentränge „R“ in der RBD-Gruppe (48,6 - 100 %), sollen an dieser Stelle

10. Zusammenfassung und Diskussion

die Leistungswerte einiger der RBD-Patienten genauer betrachtet werden (vgl. auch Tabelle 10.1).

- **Proband 12** (w, 58;3 Jahre, RBD): die Leistung von Proband 12 im Bereich der visuellen Mustererkennung liegt bei einem Prozentrang von 48,6, während der durchschnittliche Prozentrang der RBD-Gruppe bei dieser Leistung bei 73,54 liegt. Diese beiden Werte unterscheiden sich sehr signifikant voneinander ($p = .002$). Bei der auditiven Mustererkennung hingegen liegt der durchschnittliche *percent correct*-Wert der RBD-Gruppe bei 61,11 %. Die Leistung von Proband 12 mit 87,5 % ist somit signifikant höher als die der RBD-Gruppe ($p = .004$).
- **Proband 15** (m, 67;9 Jahre, RBD): im Bereich der visuellen Mustererkennung erreicht Proband 15 einen Prozentrang (AKT PR „R“) von 55,1. Dieser Wert unterscheidet sich signifikant von dem durchschnittlichen Prozentrang der RBD-Gruppe ($p = .011$). Genau umgekehrt verhält es sich bei seinen Leistungen im Bereich der auditiven Mustererkennung. Hier erreicht Proband 15 einen Wert von 100 % (*percent correct*), während der durchschnittliche Wert der RBD-Gruppe bei 61,11 % liegt ($p < .001$).
- **Proband 20** (m, 81;4 Jahre, RBD): ein umgekehrtes Bild zeigt Proband 20. Seine Leistungen bei der visuellen Mustererkennung sind tendenziell besser als die der RBD-Gruppe ($p = .073$), während sich seine Leistungen bei der Stimmerkennung sehr signifikant von dem Durchschnittswert der RBD-Patienten unterscheidet ($p = .008$).

| Proband | RBD | 12 | 15 | 20 |
|---------------------|-------|------|-------|------|
| aud. Erkennen (%) | 61,11 | 87,5 | 100 | 37,5 |
| Signifikanz (p) | - | .004 | <.001 | .008 |
| vis. Erkennen (%) | 73,54 | 48,6 | 55,1 | 85,1 |
| Signifikanz (p) | - | .002 | .011 | .073 |

Tabelle 10.1.: Einzelfalleleistungen im Vergleich zur RBD-Gesamtgruppe

10. Zusammenfassung und Diskussion

Probanden 12 und 15 zeigen also deutliche Defizite im Bereich der *visuellen* und zugleich überdurchschnittliche Leistungen im Bereich der *auditiven* Mustererkennung, während Proband 20 tendenziell bessere Leistungen bei der *visuellen* Mustererkennung erreicht, gleichzeitig aber eine stark beeinträchtigte Stimmerkennung aufweist. Diese Einzelfallbetrachtungen sollen verdeutlichen, dass zwar auf der einen Seite (mit Blick auf die Gesamtpopulation) Defizite in der visuellen und auditiven Mustererkennung bei RBD-Patienten parallel auftreten können, ein Zusammenhang zwischen diesen beiden Auffälligkeiten jedoch nicht zwingend ist. Die Vermutung, dass ein angenommener 'Metaprozess' der Mustererkennung unabhängig von der sensorischen Qualität (visuell vs. auditiv) bei einer Läsion der rechten Hemisphäre *per se* beeinträchtigt ist, muss demnach bezweifelt werden.

Die vierte Hypothese in Kapitel 8.2 besagt, dass alle Probanden unabhängig ihrer Gruppenzugehörigkeit von einer schrittweisen Verlängerung der Stimuli profitieren sollten. Dies bestätigen zahlreiche frühere Untersuchungen, u.a. von Schweinberger, Herholz & Sommer 1997, Orchard & Yarmey 1995, Yarmey & Matthys 1992, Bricker & Pruzansky 1966). Abbildung 10.3 zeigt die durchschnittliche zeitliche Dauer der verwendeten Stimuli, während Abbildung 10.4 einen Überblick über die entsprechende durchschnittliche Erkennungsleistung gibt. Die Ergebnisse früherer Untersuchungen können mit den vorliegenden Daten also bestätigt werden.

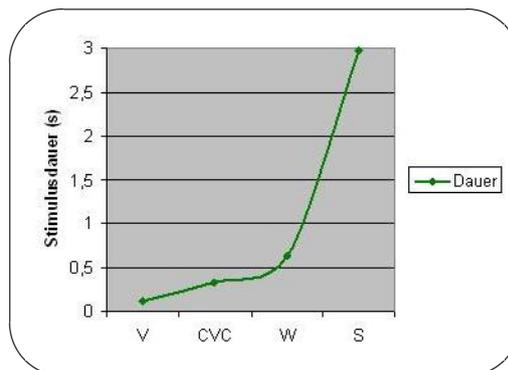


Abbildung 10.3.: Durchschnittliche zeitliche Dauer der Stimuli

Es muss jedoch angemerkt werden, dass die Leistungssteigerungen zwischen den einzelnen Items (V—CVC, CVC—W, W—S) in nur wenigen Fällen auch signifikant sind (vgl. Abbildung 10.5 auf der nächsten Seite). So lassen sich bei der LBD-Gruppe

10. Zusammenfassung und Diskussion

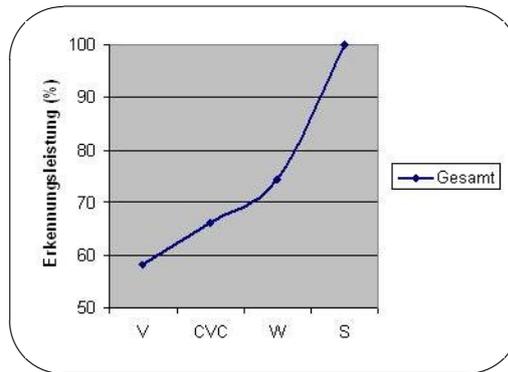


Abbildung 10.4.: Leistung der Gesamtgruppe nach Stimulidauer

keine signifikanten Leistungssteigerungen zwischen den o.g. Testitems nachweisen, dennoch steigt die Leistungskurve numerisch kontinuierlich an. Einen anderen Verlauf zeigt die Leistungskurve der RBD-Probanden. Sie beschreibt nach einem Anstieg zwischen den Items V und CVC einen Einbruch auf W-Ebene, bevor sie wieder signifikant ansteigt ($p = .003$). Die Leistungen auf W-Ebene unterscheiden sich jedoch nicht signifikant von denen auf CVC-Ebene ($p = .397$), so dass es sich auch hier lediglich um einen numerischen Unterschied und nicht um einen signifikanten Leistungseinbruch handelt.

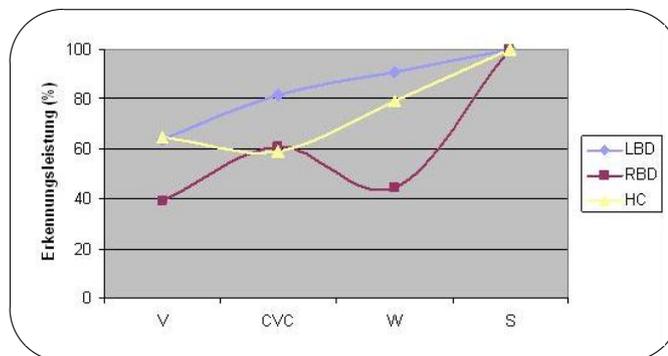


Abbildung 10.5.: Leistungsverlauf nach Stimulusdauer

Auch die Leistungskurve der HC-Probanden beschreibt einen kontinuierlichen Anstieg, obgleich sie auf CVC-Ebene kurz einbricht. Der Leistungsunterschied zwischen den Items V und CVC ist jedoch nicht signifikant ($p = .496$). Im Anschluss erreichen die HC-Probanden dann eine signifikante bzw. sehr signifikante Leistungssteigerung zwischen den Items CVC—W bzw. W—S ($p = .014$ bzw. $p = .004$).

10. Zusammenfassung und Diskussion

Während also die Leistungen der HC-Probanden auf V- bzw. CVC-Ebene statistisch gleich sind, scheinen die gesunden Kontrollprobanden dieser Untersuchung in besonderem Maße davon zu profitieren, wenn die Stimuli eine zeitliche Dauer von ≥ 500 ms aufweisen. Dies schlägt sich in dem signifikanten Anstieg der Erkennungsleistung zwischen der CVC- und der W-Ebene nieder ($p = .014$). Welche Stimulusdauer gesunde Kontrollprobanden generell benötigen, um eine Stimme korrekt zu erkennen, wird in der Literatur kontrovers diskutiert (vgl. u.a. Schweinberger, Herholz & Sommer 1997, Orchard & Yarmey 1995, Yarmey & Matthys 1992, Bricker & Pruzansky 1966; siehe dazu auch Kapitel 5.1.1). Aufgrund der großen Unterschiede zwischen den verwendeten Stimuli der verschiedenen Untersuchungen ist die Vergleich- und Generalisierbarkeit der Ergebnisse leider nicht möglich.

Ob bei der oben geschilderten Leistungssteigerung der HC-Probanden auch der semantische Inhalt der Stimuli eine Rolle gespielt hat, kann nicht abschließend beantwortet werden, da auch bisherige Untersuchungen nur wenige Hinweise auf einen etwaigen Zusammenhang geben. Die Ergebnisse der bereits im ersten Teil der Arbeit genannten Studien von Goggin et al. (1991), Thompson (1987) und Goldstein et al. (1981) legen jedoch nahe, dass gesunde Kontrollprobanden Stimmen schlechter erkennen, wenn sie in einer fremden Sprache gesprochen wurden. Auch in diesen Fällen geht der semantische Inhalt für einen Angehörigen einer anderen Sprachgemeinschaft verloren.

Bei der RBD-Gruppe zeigt sich ein anderes Bild. Da sich keine signifikanten Unterschiede zwischen den Testitems V, CVC und W zeigen, können die Erkennungsleistungen als statistisch gleich angenommen werden. Ein sprunghafter, sehr signifikanter Anstieg zeigt sich bei dieser Gruppe erst ab der W-Ebene ($p = .003$). Daraus kann geschlossen werden, dass die Patienten mit einer rechtshemisphärischen Läsion offensichtlich einen zeitlich wesentlich größeren Ausschnitt (≥ 2600 ms) einer Stimme benötigen, um ihr eine Sprecheridentität sicher zuzuweisen.

Zusammenfassend lässt sich festhalten, dass die Hypothese über eine Beeinträchtigung im Bereich der Stimmerkennung bei zerebral läsierten RBD-Patienten im Wesentlichen bestätigt werden konnte, während die Erkennungsleistung der aphasischen LBD-Patienten statistisch auf dem Niveau der Leistung der zerebral gesunden HC-Probanden liegt. Die Hypothesen zur defizitären visuellen Mustererkennungs-fähigkeit bei RBD-Patienten sowie die Abhängigkeit der Erkennungsleistung vertrauter Stimmen von deren zeitlicher Dauer können durch die Ergebnisse der vorliegenden

10. Zusammenfassung und Diskussion

Untersuchung ebenfalls unterstützt werden. Eine mit einer Hirnschädigung grundsätzlich einhergehende defizitäre Stimmerkennung (vgl. Hypothese 1) konnte jedoch nicht nachgewiesen werden.

11. Ausblick

11. Ausblick

Die vorliegende Studie zeigt, wie zahlreiche andere Untersuchungen vor ihr, dass der Prozess der auditiven Muster- bzw. Stimmerkennung bisher nicht vollständig erklärt werden kann. Unregelmäßigkeiten, Ausnahmen und Besonderheiten gerade im Bereich der Stimmerkennung lassen keine allgemein gültigen Voraussagen und Gesetzmäßigkeiten zu. Der kognitive Prozess der Stimmerkennung ist zu komplex, als dass sich aus den vorherrschenden Erklärungsmodellen zufrieden stellende Antworten und Lösungen formulieren lassen. So ist beispielsweise nach wie vor vollkommen unklar, wie sich die erhebliche Normvarianz der zerebral gesunden Kontrollprobanden erklären lässt.

Um die Fähigkeit der Stimmerkennung aphasischer und nicht-aphasischer Patienten und die diesem Prozess zugrunde liegenden zerebralen Vorgänge weiter erforschen und beleuchten zu können, stellen sich zahlreiche Fragen, deren Beantwortung einen weiteren Beitrag zum Verständnis der kognitiven Prozesse der auditiven Mustererkennung liefern können.

Eine der drängendsten Fragen ist die Überprüfung der bereits im vorigen Kapitel diskutierten Vermutung, dass die RBD-Patienten die Stimme ihrer Ehefrau nicht erkannt, sondern nach einem Ausschlussverfahren ausgewählt haben. Um diese Strategie zu umgehen wäre es nötig, die Zielstimme nicht neben unbekanntem Distraktorstimmen zu präsentieren, sondern neben weiteren vertrauten Stimmen. Diese Stimmen könnten entweder von medial bekannten Persönlichkeiten oder aber von anderen sehr vertrauten Familienmitgliedern stammen. Eine Auswahl über eine bekannt / unbekannt-Entscheidung auf Ebene der *PINs* wäre somit ausgeschlossen und Aussagen über die tatsächliche Erkennungsleistung möglich. Darüber hinaus ist es denkbar, die emotionale Reaktion des Patienten auf ein vertrautes Stimmmuster, z.B. über das Messen der Herzfrequenz, des Blutdrucks, der Atmung, der Hautleitfähigkeit oder Hormon- / Transmitterausschüttung zu erfassen. Diese Versuchsanordnung erlaubt Aussagen darüber, ob der Patient z.B. die emotional bedeutendste Stimme auswählt und dies nutzt, um seine Defizite im Bereich der Stimmerkennung zu kompensieren.

Wenn sich eine beeinträchtigte Stimmerkennungsfähigkeit bei RBD-Patienten zeigt, so kann in einem weiteren Schritt durch das Definieren mehrerer Testzeitpunkte festgestellt werden, ob die individuelle Leistungsfähigkeit unverändert bleibt, sie sich im Laufe der Zeit verändert oder durch systematische therapeutische Maßnahmen (z.B. durch die Förderung auditiver [Diskriminations-] Leistungen) beeinflussen lässt.

11. Ausblick

Eine weitere Forschungsfrage beschäftigt sich mit der Qualität der Distraktorstimmen. Wählt ein Patient / Proband nicht die Zielstimme, sondern eine der Distraktorstimmen, so können mit Hilfe einer phonetischen Analyse akustische Parameter identifiziert werden, die bei der Zielstimme und der fälschlich ausgewählten Distraktorstimme ähnlich sind. Dies ermöglicht eine Vorhersage, welche Zielstimme mit welcher Distraktorstimme verwechselt werden könnte. Das Identifizieren der akustischen Parameter im Schallsignal erlaubt Aussagen darüber, welche physikalischen Merkmale unterschiedliche Stimmen ähnlich erscheinen lassen. In einem weiteren Schritt könnte das hier vorgestellte Testdesign dadurch abgeändert werden, in dem die Distraktorstimmensets aus einer Zielstimme und akustisch / phonetisch ähnlichen Stimmen zusammengestellt würden, was eine Erschwernis des Erkennens zur Folge haben sollte.

Um den Einfluss der Stimulusdauer auf die Erkennungsleistung zu überprüfen, ist es sinnvoll, die Dauer der Stimuli möglichst kleinschrittig zu gestalten. So bietet es sich in Anlehnung an die hier verwendeten Stimuli insbesondere an, die zeitliche Länge der W-Items auf der einen und der S-Items auf der anderen Seite durch zusätzliche Stimuli zu variieren (z.B. längeres Wort – Zweiwortsatz – Dreiwortsatz etc.).

Ob dabei auch der semantische Inhalt einen Einfluss auf die Erkennungsfähigkeit zeigt, kann mit Hilfe von Nonsense-, Pseudosprach- oder Fremdsprach-Stimuli überprüft werden.

Auch die Rolle der Lautqualität kann noch differenzierter betrachtet werden, in dem beispielsweise auf CVC-Ebene Stimuli verwandt werden, die sich lediglich in einem Phonem unterscheiden (z.B. [ˈtan] vs. [ˈtɔn] vs. [ˈtʊn]).

Ein Faktor, der bisher in keiner der veröffentlichten Untersuchungen ausreichend berücksichtigt wurde, ist die Überprüfung des episodischen Gedächtnisses, in dem vermutlich die *unique pattern* der vertrauten Stimmen verortet sind. Es ist also denkbar, dass die phonagnostischen Patienten zwar einen an sich unbeeinträchtigten Analyseprozess durchlaufen, jedoch der Zugriff auf die spezifischen Informationen des episodischen Gedächtnisses unterbrochen oder beeinträchtigt sind. Zur Überprüfung bietet sich z.B. das Autobiographical Memory Interview (AMI) von Kopelman, Wilson & Baddeley (1990), der Computerisierte Gedächtnis- und Aufmerksamkeits-test – München (CGT-M) von Satzger & Engel (²1996) oder auch das Inventar zur Gedächtnisdiagnostik (IGD) von Baller et al. (2006) an.

11. Ausblick

Um genauere Erkenntnisse über den (un-) beeinträchtigten Analyseprozess der auditiven Mustererkennung zu gewinnen, ist die Entwicklung weiterer Testdesigns erforderlich, die Rückschlüsse auf die einzelnen Verarbeitungsstufen zulassen. So stellt sich die Frage, ob die hier vorgestellte Phonagnosie mit anderen auditiven Agnosien korreliert. Denkbar wären dabei z.B. Korrelationen mit der reinen Geräuschagnosie, der affektiven auditiven Agnosie, aber auch mit der Aprosodie bzw. Amusie.

Welche neuroanatomischen Korrelate bei aphasischen und nicht-aphasischen Patienten während der Stimmerkennung Aktivität zeigen, wird derzeit im Rahmen eines Kooperationsprojektes mit der Martin-Luther-Universität Halle-Wittenberg untersucht. Dazu wurde das hier verwendete Testdesign adaptiert, um die Fragestellung im fMRT untersuchen zu können (vgl. Gaul et al. [in Vorb.]).

Weitere Fragestellungen, die insbesondere durch den Einsatz bildgebender Verfahren wie PET oder fMRT beantwortet werden können, betrifft die zerebrale Aktivität bei anderen Signalqualitäten. Dieser Frage liegt die Vermutung zugrunde, dass sich das Erkennen von Stimmen von dem Erkennen anderer (Sinn tragender) Geräusche unterscheidet, Stimmen also eine gewisse „Besonderheit“ aufweisen. Es wäre also vorstellbar, dass beispielsweise ein Ornithologe bei dem Erkennen spezifischer Vogelstimmen eine ähnliche zerebrale Aktivität aufzeigt wie bei dem Erkennen von Stimmen, da er auch bei dem Erkennen von Vogelstimmen darauf angewiesen ist, relevante akustische Parameter zu extrahieren und mit bereits abgespeicherten *unique pattern* zu vergleichen. Darüber hinaus ist – wie oben bereits angedeutet – zu überprüfen, welche Rolle die Emotionalität bei der Stimmerkennung spielt¹.

Auch die bereits weiter oben vorgestellte Frage, ob Sinn tragende Stimuli andere „Erkennungsaktivitäten“ zeigen als Nonsense- oder Fremdsprachstimuli, oder ob sich im Rahmen einer Langzeitstudie (mit oder ohne therapeutische Intervention) Hirnregionen ausbilden, die die Stimmerkennungsleistung im Sinne zerebraler Reorganisation bzw. Regeneration verändern, ließe sich mit Hilfe bildgebender Verfahren beleuchten.

¹Siehe dazu auch die Studie von Alfredson et al. 2004 zur Verarbeitung emotionaler Musik.

Teil III.

Literaturverzeichnis

- Abberton, E. & Fourcin, A.J. (1978). Intonation and speaker identification. *Language and Speech*, 21:305-318.
- Abitbol, J., Abitbol, P. & Abitbol, B. (1999). Sex hormones and the female voice. *Journal of Voice*, 13:424-446.
- Akcam, T., Bolu, E., Merati, A.L., Durmus, C., Gerek, M. & Ozkaptan, Y. (2004). Voice changes after androgen therapy for hypogonadotropic hypogonadism. *The Laryngoscope*, 114:1587-1591.
- Alter, K. (2002). Suprasegmentale Merkmale und Prosodie. In H.M. Müller (Hrsg.). *Arbeitsbuch Linguistik*. Paderborn: Ferdinand Schöningh, pp. 148-169.
- Alfredson, B.B., Risberg, J., Hagberg, B. & Gustafson, L. (2004). Right temporal lobe activation when listening to emotionally significant music. *Applied Neuropsychology*, 11:161-166.
- Amir, O. & Biron-Shental, T. (2004). The impact of hormonal fluctuations on female vocal folds. *Current opinion in otolaryngology & head and neck surgery*, 12:180-184.
- Assal, G., Aubert, C. & Buttet, J. (1981). Asymétrie cérébrale et reconnaissance de la voix. *Revue Neurologique*, 137:255-268.
- Assal, G., Zander, E., Kremin, H. & Buttet J. (1976). Discrimination des voix lors des lésions du cortex cérébral. *Archives Suisses de Neurologie, Neurochirurgie et de Psychiatrie*, 119:307-315.
- Atkinson, J.E. (1976). Inter- and intraspeaker variability in fundamental voice frequency. *The Journal of the Acoustical Society of America*, 60:440-445.
- Ayotte, J., Peretz, I., Rousseau, I., Bard, C. & Bojanowski, M. (2000). Patterns of music agnosia associated with middle cerebral artery infarcts. *Brain : A Journal of Neurology*, 123:1926-1938.
- Bähr, M. & Frotscher, M. (2003). *Duus' Neurologisch-topische Diagnostik*. Stuttgart, New York: Thieme.
- Baldwin J. & French, P. (Eds.) (1990). *Forensic Phonetics*. London: Pinter.
- Baller, G., Brand, M., Kalbe, E. & Kessler, J. (2006). *Inventar zur Gedächtnisdiagnostik*. Hogrefe: Göttingen.

- Bartholomeus, B. (1973). Voice identification by nursery school children. *Canadian Journal of Psychology*, 27:464-472.
- Bauer, R.M. & Demery, J.A. (2003). Agnosia. In K.M. Heilman & E. Valenstein (Eds.), *Clinical Neuropsychology*. Oxford: Oxford University Press, pp. 236-295.
- Bautista, R.E. & Ciampetti, M.Z. (2003). Expressive aprosody and amusia as a manifestation of right hemisphere seizures. *Epilepsia*, 44:466-467.
- Belin, P., Fecteau, S. & Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences*, 8:129-135.
- Belin, P. & Zatorre, R.J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport*, 14:2105-2109.
- Belin, P., Zatorre, R.J. & Ahad P. (2002). Human temporal-lobe response to vocal sounds. *Cognitive Brain Research*, 13:17-26.
- Bennett, S. & Montero-Diaz, L. (1982). Children's perception of speaker sex. *Journal of Phonetics*, 10:113-121.
- Berendes, J. (1962). Veränderungen der weiblichen Stimme durch virilisierende und anabole Hormone. *Folia Phoniatria et Logopaedica*, 14:265-271.
- Binieck, R. (1997). *Akute Aphasien*. Stuttgart, New York: Thieme.
- Böhme, G. & Hecker, G. (1970). Gerontologische Untersuchungen über Stimmumfang und Sprechstimmlage. *Folia Phoniatria et Logopaedica*, 22:176-184.
- Boller, F. & Green, E. (1972). Comprehension severe aphasia. *Cortex*, 8:382-394.
- Boucher, J., Lewis, V. & Collis, G. (1998). Familiar face and voice matching and recognition in children with autism. *Journal of Child Psychology and Psychiatry, and allied Disciplines*, 39:171-181.
- Boucher, J., Lewis, V. & Collis, G.M. (2000). Voice processing abilities in children with autism, children with specific language impairments, and young typically developing children. *Journal of Child Psychology and Psychiatry, and allied Disciplines*, 41:847-857.
- Boulet, M.J. & Oddens, B.J. (1996). Female voice changes around and after the menopause – an initial investigation. *Maturitas*, 23:15-21.
- Brachacki, G.W., Fawcett, A.J. & Nicolson, R.I. (1994). Adults with dyslexia have a deficit in voice recognition. *Perceptual and Motor Skills*, 78:304-306.

- Braun, A. (1995). Fundamental frequency – how speaker-specific is it? In A. Braun & J.-P. Köster (Eds.). *Studies in Forensic Phonetics*. Trier: Wissenschaftlicher Verlag, pp. 9-23.
- Breznitz, Z. (1992). Verbal indicators of depression. *The Journal of General Psychology*, 119:351-363.
- Bricker, P.D. & Pruzansky, S. (1966). Effects of stimulus content and duration on talker identification. *The Journal of the Acoustical Society of America*, 40:1441-1449.
- Bricker, P.D. & Pruzansky, S. (1974). Speaker Recognition. In N. Lass (Ed.). *Experimental Phonetics*. New York: MSS Information Corp. pp. 295-326.
- Broeders, A.P.A., Cambier-Langeveld, T. & Vermeulen, J. (2002). Arranging a voice line-up in a foreign language. *Forensic Linguistics*, 9:104-112.
- Broeders, A.P.A. & Rietveld, A.C.M. (1995). Speaker Identification by Earwitnesses. In A. Braun & J.-P. Köster (Eds.). *Studies in Forensic Phonetics*. Trier: Wissenschaftlicher Verlag, pp. 24-40.
- Brown, R. (1981). An experimental study of the relative importance of acoustic parameters for auditory speaker recognition. *Language and Speech*, 24:295-310.
- Bruce, V. & Valentine, T. (1985). Identity priming in the recognition of familiar faces. *British Journal of Psychology*, 76:373-383.
- Bruce, V. & Young, A. (1986). Understanding Face Recognition. *British Journal of Psychology*, 77:305-327.
- Bull, R. & Clifford, B.R. (1987). Earwitness voice recognition accuracy. In G.L. Wells & E.F. Loftus (Eds.), *Eyewitness Testimony: Psychological Perspectives*. New York: Cambridge University Press, pp. 92-123.
- Bull, R., Rathborn, H. & Clifford, B.R. (1983). The voice-recognition accuracy of blind listeners. *Perception*, 12:223-226.
- Burton, A.M., Bruce, V. & Johnston, R.A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology*, 81:361-380.
- Cerrato, L., Falcone, M. & Paoloni, A. (2000). Subjective age estimation of telephonic voices. *Speech Communication*, 31:107-112.

- Charrier, I., Mathevon, N. & Jouventin, P. (2002). How does a fur seal mother recognize the voice of her pup? An experimental study of *Arctocephalus tropicalis*. *The Journal of Experimental Biology*, 205:603-612.
- Charrier, I., Mathevon, N. & Jouventin, P. (2001). Mother's voice recognition by seal pups. *Nature*, 412:873.
- Clarke, F.R. & Becker, R.W. (1969). Comparison of techniques for discriminating among talkers. *Journal of Speech and Hearing Research*, 12:747-761.
- Clifford, B.R., Rathborn, H. & Bull, R. (1981). The effects of delay on voice recognition accuracy. *Law and Human Behaviour*, 5:201-208.
- Colton, R.H. & Estill, J. (1979). Perceptual aspects of some voice qualities. *The Journal of the Acoustical Society of America*, 65 (Supplement 1):115.
- Compton, A.J. (1963). Effects of filtering and vocal duration upon the identification of speakers, aurally. *The Journal of the Acoustical Society of America*, 35:1748-1752.
- Confavreux, C., Croisile, B., Garassus, P., Aimard, G. & Trillet, M. (1998). Progressive amusia and aprosody. *Archives of Neurology*, 49:971-976.
- Cosentino, S. & Stern, Y. (2005). Metacognitive theory and assessment in dementia: do we recognize our areas of weakness? *Journal of the International Neuropsychological Society*, 11:910-919.
- Coslett, H.B., Brashear, H.R. & Heilman, K.M. (1984). Pure word deafness after bilateral primary auditory cortex infarcts. *Neurology*, 34:347-352.
- Covington, M.A., He, C., Brown, C., Naci, L., McClain, J.T., Fjordbak, B.S., Semple, J. & Brown, J. (2005). Schizophrenia and the structure of language: the linguist's view. *Schizophrenia Research*, 77:85-98.
- Crinion, J.T., Lambon-Ralph, M.A., Warburton, E.A., Howard, D. & Wise, R.J. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain: A Journal of Neurology*, 126:1193-1201.
- Cutler, B.L. & Penrod, S.D. (1990). Faktoren, die die Zuverlässigkeit der Personenidentifizierung beeinflussen. In G. Köhnken & S.L. Sporer (Hrsg.), *Identifizierung von Tatverdächtigen durch Augenzeugen*. Stuttgart: Verlag für Angewandte Psychologie, pp. 25-51.

- DeCasper, A.J. & Fifer, W.P. (1980). Of human bonding: newborns prefer their mothers' voices. *Science*, 208:1174-1176.
- De Renzi, E. & Vignolo L.A. (1962). The Token Test: a sensitive test to detect receptive disturbances in aphasics. *Brain: A Journal of Neurology*, 85:665-678.
- Doty, N. (1998). The influence of nationality on the accuracy of face and voice recognition. *The American Journal of Psychology*, 111:191-214.
- Duffy, R.J. (1970). Fundamental frequency characteristics of adolescent females. *Language and Speech*, 13:14-24.
- Ellis, A.W., Young, A.W. & Critchley, E.M. (1989). Loss of memory for people following temporal lobe damage. *Brain: A Journal of Neurology*, 112:1469-1483.
- Ellis HD, Young AW, Quayle AH, De Pauw KW. (1997). Reduced autonomic responses to faces in Capgras delusion. *Proceedings - Biological Sciences / The Royal Society*, 264:1085-1092.
- Endres, W., Bambach, W. & Flösser, G. (1971). Voice spectrograms as a function of age, voice disguise, and voice imitation. *The Journal of the Acoustical Society of America*, 49:1842-1848.
- Engelien, A. (2003). Auditive Agnosien. In H.-O. Karnath & P. Thier (Hrsg.), *Neuropsychologie*. Berlin, Heidelberg, New York: Springer-Verlag, pp. 139-144.
- Engelien, A. (2002). Störungen des auditiven Verarbeitens. In W. Hartje & K. Poeck (Hrsg.), *Klinische Neuropsychologie*. Stuttgart, New York: Thieme, pp. 435-442.
- Engelien, A., Huber, W., Silbersweig, D., Stern, E., Frith, C.D., Doring, W., Thron, A. & Frackowiak, R.S. (2000). The neural correlates of 'deaf-hearing' in man: conscious sensory awareness enabled by attentional modulation. *Brain: A Journal of Neurology*, 123:532-545.
- Engelien, A., Silbersweig, D., Stern, E., Huber, W., Doring, W., Frith, C. & Frackowiak, R.S. (1995). The functional anatomy of recovery from auditory agnosia. A PET study of sound categorization in a neurological patient and normal controls. *Brain: A Journal of Neurology*, 118:1395-1409.
- Erickson, D., Yoshida, K., Menezes, C., Fujino, A., Mochida, T. & Shibuya, Y. (2006). Exploratory study of some acoustic and articulatory characteristics of sad speech. *Phonetica*, 63:1-25.

- Faller, A. & Schünke, M. (¹⁴2004). *Der Körper des Menschen – Einführung in Bau und Funktion*. Stuttgart, New York: Thieme.
- Fant, G. (1960). *Acoustic Theory of Speech Production*. Den Haag, Paris: Mouton.
- Fant, G. (1974). *Speech Sounds and Features*. Cambridge (Mass.), London: The MIT Press.
- Ferrand, C.T. (2002). Harmonics-to-noise ratio: an index of vocal aging. *Journal of Voice*, 16:480-487.
- Fitch, W.T. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences*, 4:258-267.
- Fujii, T., Fukatsu, R., Watabe, S., Ohnuma, A., Teramura, K., Kimura, I., Saso, S. & Kogure, K. (1990). Auditory sound agnosia without aphasia following a right temporal lobe lesion. *Cortex*, 26:263-268.
- Gatterer, G. (1990). *Alters-Konzentrations-Test*. Göttingen: Hogrefe.
- Gaul, C., Kneidl, O., Knörger, M., Lehmann, N., Paelecke-Habermann, Y. & Schönfeld, R. (*in Vorb.*). Voice Perception by Aphasics: an fMRI-Study.
- George, M.S., Parekh, P.I., Rosinsky, N., Ketter, T.A., Kimbrell, T.A., Heilman, K.M., Herscovitch, P. & Post, R.M. (1996). Understanding emotional prosody activates right hemisphere regions. *Archives of Neurology*, 53:665-670.
- Goggin, J.P., Thompson, C.P., Strube, G. & Simental, L.R. (1991). The role of language familiarity in voice identification. *Memory and Cognition*, 5:448-458.
- Goldenberg, G. (⁵2002a). Störungen des visuellen Erkennens und des bildlichen Vorstellens. In W. Hartje & K. Poeck (Hrsg.), *Klinische Neuropsychologie*. Stuttgart, New York: Thieme, pp. 296-315.
- Goldenberg, G. (³2002b). *Neuropsychologie*. München, Jena: Urban & Fischer.
- Goldman-Eisler, F. (²1973). *Psycholinguistics. Experiments in Spontaneous Speech*. London, New York: Academic Press.
- Goldstein, A.G., Knight, P., Bailis, K. & Conover, J. (1981). Recognition memory for accented and unaccented voices. *Bulletin of the Psychonomic Society*, 17:217-220.
- Gonzalez J. & Carpi A. (2004). Early effects of smoking on the voice: a multidimensional study. *Medical Science Monitor*, 10:649-656.

- Goodglass, H., Fordor, I. & Schulhoff, C. (1967). Prosodic factors in Grammar – Evidence from aphasia. *Journal of Speech and Hearing Research*, 10:5-20.
- Gorham-Rowan M.M. & Laures-Gore J. (2006). Acoustic-perceptual correlates of voice quality in elderly men and women. *Journal of Communication Disorders*, 39:171-184.
- Grabe, E. & Nolan, F. (1996). Preparing a voice lineup. *Forensic Linguistics*, 3:74-94.
- Green, D.M. & Swets, J.A. (³1988). *Signal Detection Theory and Psychophysics*. New York: R.E. Krieger Co.
- Günzburger, D. (1995). Acoustic and perceptual implications of the transsexual voice. *Archives of Sexual Behavior*, 24:339-348.
- Günzburger, D., Bresser, A. & ter Keurs, M. (1987). Voice identification of prepubertal boys and girls by normally sighted and visually handicapped subjects. *Language and Speech*, 30:47-58.
- Guimaraes, I. & Abberton, E. (2005). Health and voice quality in smokers: an exploratory investigation. *Logopedics, Phoniatrics, Vocology*, 30:185-91.
- Härting, C., Markowitsch, H.J., Neufeld, H., Calabrese, P. & Deisinger, K. (Hrsg.) (2000). *Wechsler Gedächtnis Test - Revidierte Fassung (WMS-R)*. Göttingen: Hogrefe.
- Hall, D.A., Johnsrude, I.S., Haggard, M.P., Palmer, A.R., Akeroyd, M.A. & Summerfield, A.Q. (2002). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 12:140-149.
- Hargreaves, W.A. & Starkweather, J.A. (1963). Recognition of speaker identity. *Language and Speech*, 6:63-67.
- Hart, H.C., Palmer, A.R. & Hall, D.A. (2002). Amplitude and frequency-modulated stimuli activate common regions of human auditory cortex. *Cerebral Cortex*, 13:773-781.
- Hartje, W. & Poeck, K. (⁵2002). *Klinische Neuropsychologie*. Stuttgart, New York: Thieme.
- Hattiangadi, N., Pillion, J.P., Slomine, B., Christensen, J., Trovato, M.K. & Speedie, L.J. (2005). Characteristics of auditory agnosia in a child with severe traumatic brain injury: a case report. *Brain and Language*, 92:12-25.

- Hecker, M.H. (1971). *Speaker Recognition – An Interpretive Study of the Literature. ASHA Monographs Number 16*. American Speech and Hearing Association: Washington D.C.
- Heckmann, J.G., Lang, C.J., Lösslein, H., von Stockert, T. & Neundörfer, B. (1994). Untersuchung zum Erkennen vertrauter Handschrift nach linkshemisphärischer und rechtshemisphärischer Hirnschädigung. *Der Nervenarzt*, 65:396-401.
- Heckmann, J.G., Lang, C.J. & Neundörfer, B. (2001). Recognition of familiar handwriting in stroke and dementia. *Neurology*, 57:2128-2131.
- Heike, G. (1990). Stimmenvergleichsgutachten – Probleme und Methoden der forensischen Phonetik am Beispiel einer Fallstudie. In H. Kniffka (Hrsg.). *Texte zu Theorie und Praxis forensischer Linguistik*. Tübingen: Max Niemeyer Verlag, pp. 481-489.
- Heilman, K.M., Scholes, R. & Watson, R.T. (1975). Auditory affective agnosia. Disturbed comprehension of affective speech. *Journal of Neurology, Neurosurgery, and Psychiatry*, 38:69-72.
- Helfrich, H. (1979). Age markers in speech. In K.R. Scherer & H. Giles (Eds.). *Social markers in Speech*. Cambridge, London, New York, Melbourne: Cambridge University Press, pp. 63-107.
- Hepper, P., Scott, D. & Shahidullah, S. (1993). Newborn and fetal response to maternal voice. *Journal of Reproductive and Infant Psychology*, 11:147-153.
- Herrmann-Lingen, C., Buss, U. & Snaith, R.P. (1995). *Hospital Anxiety and Depression Scale – Deutsche Version*. Göttingen: Hogrefe.
- Hollien, H. (1990). *The Acoustics of Crime. The New Science of Forensic Phonetics*. New York, London: Plenum-Press.
- Hollien, H., Bennett, G. & Gelfer, M.P. (1983). Criminal identification comparison: aural versus visual identifications resulting from a simulated crime. *Journal of Forensic Sciences*, 28:208-221.
- Hollien, H., Green, R., Massey, K. (1994). Longitudinal research on adolescent voice change in males. *The Journal of the Acoustical Society of America*, 96:2646-2654.

- Hollien, H., Majewski, W. & Doherty, E.T. (1982). Perceptual identification of voices under normal, stress, and disguise speaking conditions. *Journal of Phonetics*, 10:139-148.
- Hollien, H. & Paul, P. (1969). A second evaluation of the speaking fundamental frequency characteristics of post-adolescent girls. *Language and Speech*, 12:119-124.
- Hollien, H. & Shipp, T. (1972). Speaking fundamental frequency and chronologic age in males. *Journal of Speech and Hearing Research*, 15:155-159.
- Holmgren, G.L. (1967). Physical and psychological correlates of speaker recognition. *Journal of Speech and Hearing Research*, 10:57-66.
- Homae, F., Watanabe, H., Nakano, T., Asakawa, K. & Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neuroscience Research*, 54:276-280.
- Huber, W., Poeck, K. & Weniger, D. (⁵2002). Aphasie. In: W. Hartje & K. Poeck (Hrsg.). *Klinische Neuropsychologie*. Stuttgart, New York: Thieme, pp. 93-173
- Huber, W., Poeck, K., Weniger, D. & Willmes, K. (1983). *Aachener Aphasie Test*. Göttingen: Hogrefe.
- Imaizumi, S., Mori, K., Kiritani, S., Kawashima, R., Sugiura, M., Fukuda, H., Itoh, K., Kato, T., Nakamura, A., Hatano, K., Kojima, S. & Nakamura, K. (1997). Vocal identification of speaker and emotion activates different brain regions. *Neuroreport*, 8:2809-2812.
- Insley, S.J. (2000). Long-term vocal recognition in the northern fur seal. *Nature*, 406:404-405.
- Jerger, J., Weikers, N.J., Sharbrough, F.W. (3rd) & Jerger, S. (1969). Bilateral lesions of the temporal lobe. A case study. *Acta Oto-Laryngologica - Supplementum*, 258:1-51.
- Jessen, M. (2004). *Forensische Sprechererkennung in Forschung und Praxis*. Eigene Aufzeichnungen zum SFB-Vortrag am 26.01.2004 an der Universität Bielefeld.
- Jones, D. (⁹1975, ¹1918) *An outline of English phonetics*. Cambridge: Cambridge University Press.

- Johnsrude, I.S., Penhune, V.B. & Zatorre R.J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain: A Journal of Neurology*, 123:155-163.
- Johnstone, T., van Reekum, C.M., Hird, K., Kirsner, K. & Scherer, K.R. (2005). Affective speech elicited with a computer game. *Emotion*, 5:513-518.
- Jonsdottir, V., Laukkanen, A.M. & Vilkmann E. (2002). Changes in teachers' speech during a working day with and without electric sound amplification. *Folia phoniatrica et logopaedica*, 54:282-287.
- Kaga, K., Kaga, M., Tamai, F. & Shindo, M. (2003). Auditory agnosia in children after Herpes Encephalitis. *Acta Oto-Laryngologica*, 123:232-235.
- Kaga, K., Kurauchi, T., Nakamura, M., Shindo, M. & Ishii, K. (2005). Magnetoencephalography and positron emission tomography studies of a patient with auditory agnosia caused by bilateral lesions confined to the auditory radiations. *Acta Oto-Laryngologica*, 125:1351-1355.
- Kanwisher, N., McDermott, J. & Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, 17:4302-4311.
- Kapur, N., Ellison, D., Parkin, A.J., Hunkin, N.M., Burrows, E., Sampson, S.A. & Morrison, E.A. (1994). Bilateral temporal lobe pathology with sparing of medial temporal lobe structures: lesion profile and pattern of memory disorder. *Neuropsychologia*, 32:23-38.
- Kapur, N., Ellison, D., Smith, M.P., McLellan, D.L. & Burrows, E.H. (1992). Focal retrograde amnesia following bilateral temporal lobe pathology. A neuropsychological and magnetic resonance study. *Brain: A Journal of Neurology*, 115:73-85.
- Kazui, S., Naritomi, H., Sawada, T., Inoue, N. & Okuda, J. (1990). Subcortical auditory agnosia. *Brain and Language*, 38:476-487.
- Kerstholt, J.H., Jansen, N.J.M., Van Amelsvoort, A.G. & Broeders, A.P.A. (2004). Earwitnesses: effects of speech duration, retention interval and acoustic environment. *Applied Cognitive Psychology*, 18:327-336.
- Kerstholt, J.H., Jansen, N.J.M., Van Amelsvoort, A.G. & Broeders, A.P.A. (2006). Earwitnesses: effects of accent, retention and telephone. *Applied Cognitive Psychology*, 20:187-197.

- Kimelman, M.D. (1991). The role of target word stress in auditory comprehension by aphasic listeners. *Journal of Speech and Hearing Research*, 34:334-339.
- Kimelman, M.D. (1999). Prosody, linguistic demands, and auditory comprehension in aphasia. *Brain and Language*, 69:212-221.
- Kisilevsky, B.S., Hains, S.M., Lee, K., Xie, X., Huang, H., Ye, H.H., Zhang, K. & Wang Z. (2003). Effects of experience on fetal voice recognition. *Psychological Science: a Journal of the American Psychological Society*, 14:220-224.
- Klingholz, F., Penning, R. & Liebhardt, E. (1988). Recognition of low-level alcohol intoxication from speech signal. *The Journal of the Acoustical Society of America*, 84:929-935.
- Kopelmann, M., Wilson, B. & Baddeley, A. (1990). *Autobiographical Memory Interview*. Göttingen: Hogrefe.
- Krahe, B. & Kundrotas, S. (1992) Glaubwürdigkeitsbeurteilung bei Vergewaltigungsanzeigen: Ein aussagenanalytisches Feldexperiment. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 39:598-620.
- Kreiman, J., Gerratt, B.R., Kempster, G.B., Erman, A. & Berke, G.S. (1993). Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *Journal of Speech and Hearing Research*, 36:21-40.
- Kreiman, J., Gerratt, B.R. & Precoda, K. (1990). Listener experience and perception of voice quality. *Journal of Speech and Hearing Research*, 33:103-115.
- Kreiman, J., Gerratt, B.R., Precoda, K. & Berke, G.S. (1992). Individual differences in voice quality perception. *Journal of Speech and Hearing Research*, 35:512-520.
- Krook, M.I.P. (1988). Speaking fundamental frequency characteristics of normal Swedish subjects obtained by glottal frequency analysis. *Folia Phoniatrica et Logopaedica*, 40:82-90.
- Kohler, K.J. (1995). *Einführung in die Phonetik des Deutschen*. Berlin: Erich Schmidt Verlag.
- Künzel, H.J. (1987). *Sprechererkennung – Grundzüge forensischer Sprachverarbeitung*. Heidelberg: Kriminalistik-Verlag.
- Künzel, H.J. (1989). Die Erkennung von Personen anhand ihrer Stimme. *Neue Zeitschrift für Strafrecht*, 9:400-405.

- Künzel, H.J. (1990). *Phonetische Untersuchungen zur Sprecher-Erkennung durch linguistisch naive Personen*. Stuttgart: Steiner.
- Künzel, H.J. (1995). Field procedures in forensic speaker recognition. In J.W. Lewis (Ed.), *Studies in General and English Phonetics – Essays in Honour of J.D. O'Connor*. London: Routledge, pp. 68-84.
- Künzel, H.J. (1997). Methoden der forensischen Sprecher-Erkennung. *Strafverteidiger-Forum*, 9:100-105.
- Künzel, H.J. (1998). Zur Entwicklung der forensischen Sprecher-Erkennung in Deutschland. In Bundeskriminalamt (Hrsg.). *Festschrift für Horst Herold zum 75. Geburtstag*. Wiesbaden: Bundeskriminalamt, pp. 369-395.
- Künzel, H.J. (2001). Beware of the 'telephone effect': the influence of telephone-transmission on the measurement of formant frequencies. *Forensic Linguistics*, 8:80-99.
- Künzel, H.J., Braun, A. & Eysholdt, U. (1992). *Einfluss von Alkohol auf Sprache und Stimme*. Heidelberg: Kriminalistik-Verlag.
- Kuttruff, H. (2004). *Akustik*. Stuttgart, Leipzig: S. Hirzel.
- Ladefoged, P. (2006). *A Course in Phonetics*. Boston: Thomson Wadsworth.
- Ladefoged, P. & Ladefoged, J. (1980). The ability of listeners to identify voices. *UCLA Working Papers in Phonetics*, 41:43-51.
- Lambert, J., Eustache, F., Lechevalier, B., Rossa, Y. & Viader F. (1989). Auditory agnosia with relative sparing of speech perception. *Cortex*, 25:71-82.
- Langenmayr, A. (1997). *Sprachpsychologie – Ein Lehrbuch*. Göttingen, Bern, Toronto, Seattle: Hogrefe.
- Lass, N.J., Hughes, K.R., Bowyer, M.D., Waters, L.T. & Bourne, V.T. (1979). Speaker sex identification from voiced, whispered, and filtered isolated vowels. *The Journal of the Acoustical Society of America*, 59:675-678.
- Lattner, S., Meyer, M.E. & Friederici, A.D. (2005). Voice perception: sex, pitch, and the right hemisphere. *Brain Mapping*, 24:11-20.
- Laver, J. (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.

- Lavner, Y., Gath, I. & Rosenhouse, J. (2000). The effects of acoustic modifications on the identification of familiar voices speaking isolated vowels. *Speech Communication*, 30:9-26.
- Legge, G.E., Grosmann, C. & Pieper, C.M. (1984). Learning unfamiliar voices. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10:289-303.
- Leischner, A. (²1987). *Aphasien und Sprachentwicklungsstörungen*. Stuttgart, New York: Thieme.
- Liégeois-Chauvel, C., Peretz, I., Babai, M., Laguitton, V. & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain: A Journal of Neurology*, 121:1853-1867.
- Linville, S.E. (1988). Intraspeaker variability in fundamental frequency stability: an age-related phenomenon? *The Journal of the Acoustical Society of America*, 83:741-745.
- Linville, S.E. & Korabic, E.W. (1987). Fundamental frequency stability characteristics of elderly women's voices. *The Journal of the Acoustical Society of America*, 81:1196-1199.
- Mann, V.A., Diamond, R. & Carey, S. (1979). Development of voice recognition: Parallels with face recognition. *Journal of Experimental Child Psychology*, 27:153-165.
- Markowitsch, H.J. (2003). Psychogenic amnesia. *NeuroImage*, 20 (Suppl 1):S132-S138.
- Markowitsch, H.J. (1995). Which brain regions are critically involved in the retrieval of old episodic memory? *Brain Research Reviews*, 21:117-127.
- Markowitsch, H.J., Thiel, A., Reinkemeier, M., Kessler, J., Koyuncu, A. & Heiss, W.D. (2000). Right amygdalar and temporofrontal activation during autobiographic, but not during fictitious memory retrieval. *Behavioural Neurology*, 12:181-190.
- Marx, H. (1997). Erwerb des Lesens und des Rechtschreibens: Literaturüberblick. In F.E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter*. Weinheim: Psychologie Verlags Union, pp. 85-111.

- McGehee, F. (1937). The reliability of the identification of the human voice. *Journal of General Psychology*, 17:249-271.
- McGehee, F. (1944). An experimental study of voice recognition. *Journal of General Psychology*, 31:53-65.
- McGlone, R.E. & Hollien, H. (1963). Vocal pitch characteristics of old women. *Journal of Speech and Hearing Research*, 6:164-170.
- Mendez, M.F. (2001). Generalized auditory agnosia with spared music recognition in a left-hander. Analysis of a case with a right temporal stroke. *Cortex*, 37:139-150.
- Mesulam, M.M. (1998). From sensation to cognition. *Brain: a Journal of Psychology*, 121:1013-1052.
- Meurer, E.M., Wender, M.C., von Eye Corleta, H. & Capp, E. (2004). Phono-articulatory variations of women in reproductive age and postmenopausal. *Journal of Voice*, 18:369-374.
- Michel, F., Peronnet, F. & Schott, B. (1980). A case of cortical deafness: clinical and electrophysiological data. *Brain and Language*, 10:367-377.
- Mitchell, R.L., Elliott, R., Barry, M., Cruttenden, A. & Woodruff, P.W. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia*, 41:1410-1421.
- Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76:165-178.
- Motomura, N., Yamadori, A., Mori, E. & Tamaru, F. (1986). Auditory agnosia. Analysis of a case with bilateral subcortical lesions. *Brain: A Journal of Neurology*, 109:379-391.
- Murphy, C.H. & Doyle, P.C. (1987). The effects of cigarette smoking on voice-fundamental frequency. *Otolaryngology and Head and Neck Surgery*, 97:376-380.
- Mysak, E. (1959). Pitch and duration characteristics of older males. *Journal of Speech and Hearing Research*, 2:46-54.
- Nakamura, K., Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Nagumo, S., Kubota, K., Fukuda, H., Ito, K. & Kojima S. (2001). Neural

- substrates for recognition of familiar voices: a PET study. *Neuropsychologia*, 39:1047-1054.
- Nakamura, K. & Kubota, K. (1996). The primate temporal pole: its putative role in object recognition and memory. *Behavioural Brain Research*, 77:53-77.
- Neppert, J.M.H. (1999). *Elemente einer Akustischen Phonetik*. Hamburg: Helmut Buske Verlag.
- Neuner, F. & Schweinberger, S.R. (2000). Neuropsychological impairments in the recognition of faces, voices, and personal names. *Brain and Cognition*, 44:342-366.
- Nilsson, A. (1988). Speech characteristics as indicators of depressive illness. *Acta Psychiatrica Scandinavica*, 77:253-263.
- Nilsson, A. (1987). Acoustic analysis of speech variables during depression and after improvement. *Acta Psychiatrica Scandinavica*, 76:235-245.
- Nolan, F. (1990). The limitations of auditory-phonetic speaker identification. In H. Kniffka (Hrsg.). *Texte zu Theorie und Praxis forensischer Linguistik*. Tübingen: Max Niemeyer Verlag, pp. 457-479.
- Nolan, F. (1983). *The Phonetic Bases of Speaker-Recognition*. Cambridge: Cambridge University Press.
- Nygaard, L.C. & Pisoni, D.B. (1998). Talker-specific learning in speech perception. *Perception and Psychophysics*, 60:355-376.
- Orchard, T.-L., Yarmey, A.D. (1995). The effects of whispers, voice-sample duration, and voice distinctiveness on criminal speaker identification. *Applied Cognitive Psychology*, 9:249-260.
- Orgass, B., De Renzi, E. & Vignolo, L.A. (1982). *Token Test*. Göttingen: Hogrefe.
- Orimo, S., Ozawa, E., Uematsu, M., Yoshida, E., Hino, H., Yamada, M., Okeda, R. & Mizusawa, H. (2000). A case of Creutzfeldt-Jakob disease presenting with auditory agnosia as an initial manifestation. *European Neurology*, 44:256-258.
- Orlikoff, R.F. & Baken, R.J. (1989a). The effect of the heartbeat on vocal fundamental frequency perturbation. *Journal of Speech and Hearing Research*, 32:576-582.

- Orlikoff, R.F. & Baken, R.J. (1989b). Fundamental frequency modulation of the human voice by the heartbeat: preliminary results and possible mechanisms. *The Journal of the Acoustical Society of America*, 85:888-893.
- Osgood, C.E., Suci, G.J. & Tannenbaum, P.H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.
- Pan, C.L., Kuo, M.F. & Hsieh, S.T. (2004). Auditory agnosia caused by a tectal germinoma. *Neurology*, 63:2387-2389.
- Pedersen, M.F., Moller S., Krabbe S., Bennett P. & Svenstrup B. (1990). Fundamental voice frequency in female puberty measured with electroglottography during continuous speech as a secondary sex characteristic. A comparison between voice, pubertal stages, oestrogens and androgens. *International Journal of Pediatric Otorhinolaryngology*, 20:17-24.
- Pedersen, M.F., Moller S., Krabbe, S. & Bennett, P. (1986). Fundamental voice frequency measured by electroglottography during continuous speech. A new exact secondary sex characteristic in boys in puberty. *International Journal of Pediatric Otorhinolaryngology*, 11:21-27.
- Pell, M.D. (2006). Cerebral mechanisms for understanding emotional prosody in speech. *Brain and Language*, 96:221-234.
- Pell, M.D. (1999). Fundamental frequency encoding of linguistic and emotional prosody by right hemisphere-damaged speakers. *Brain and Language*, 69:161-192.
- Pell, M.D. (1998). Recognition of prosody following unilateral brain lesion: influence of functional and structural attributes of prosodic contours. *Neuropsychologia*, 36:701-715.
- Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain: A Journal of Neurology*, 113:1185-1205.
- Peretz, I. (2001). Music perception and recognition. In B. Rapp (Ed.). *The Handbook of Cognitive Neuropsychology*. Hove, UK: Psychology Press. pp. 519-540.
- Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G. & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain: A Journal of Neurology*, 117:1283-1301.

- Peretz, I. & Zatorre, R.J. (2004). Brain organization for music processing. *Annual Review of Psychology*, 56:89-114.
- Perry, T.L., Ohde, R.N. & Ashmead, D.H. (2001). The acoustic bases for gender identification from children's voices. *The Journal of the Acoustical Society of America*, 109:2988-2998.
- Pétursson, M. & Neppert, J.M.H. (³2002). *Elementarbuch der Phonetik*. Hamburg: Helmut Buske Verlag.
- Peuser, G. & Winter, S. (2000). *Lexikon zur Sprachtherapie: Terminologie der Patholinguistik*. München: Fink.
- Pisoni, D.B. & Martin, C.S. (1989). Effects of alcohol on the acoustic-phonetic properties of speech: perceptual and acoustic analyses. *Alcoholism: Clinical and Experimental Research*, 13:577-587.
- Pollack, I., Pickett, J.M. & Sumbly, W.H. (1954). On the identification of speakers by voice. *The Journal of the Acoustical Society of America*, 26:403:406.
- Pompino-Marschall, B. (²2003). *Einführung in die Phonetik*. Berlin, New York: de Gruyter.
- Protopapas, A. & Lieberman, P. (1997). Fundamental frequency of phonation and perceived emotional stress. *The Journal of the Acoustical Society of America*, 101:2267-2277.
- Pschyrembel Klinisches Wörterbuch (²⁵⁹2002). Berlin: de Gruyter.
- Ptacek, P.H. & Sander, E.K. (1966). Age recognition from voice. *Journal of Speech and Hearing Research*, 9:273-277.
- Pullum, G.K. & Ladusaw, W.A. (²1996). *Phonetic Symbol Guide*. Chicago, London: University of Chicago Press.
- Putz, R. & Pabst, R. (Hrsg.) (²¹2004). *Sobotta – Atlas der Anatomie des Menschen*. München, Jena: Urban & Fischer.
- Raithel, V. (2005). *The Perception of Intonation Contours and Focus by Aphasic and Healthy Individuals*. Tübingen: Günter Narr Verlag.
- Rantala, L., Vilkmann, E. & Bloigu, R. (2002). Voice changes during work: subjective complaints and objective measurements for female primary and secondary schoolteachers. *Journal of Voice*, 16:344-355.

- Roe, K.V. (1978). Mother-stranger discrimination in three-month-old infants and subsequent Gesell performance. *The Journal of Genetic Psychology; Child Behavior, Animal Behavior, and Comparative Psychology*, 133:111-118.
- Roebuck, R. & Wilding, J. (1993). Effects of vowel variety and sample length on identification of a speaker in a line-up. *Applied Cognitive Psychology*, 7:475-481.
- Rose, P. (1999a). Differences and distinguishability in the acoustic characteristics of *Hello* in voices of similar-sounding speakers. *Australian Review of Applied Linguistics*, 21:1-42.
- Rose, P. (1999b). Long- and short-term within-speaker differences in the formants of Australian 'Hello'. *Journal of the International Phonetic Association*, 29:1-31.
- Rose, P. (2002). *Forensic Speaker Identification*. London, New York: Taylor & Francis.
- Satzger, W. & Engel, R.R. (1996). *Computerisierter Gedächtnis- und Aufmerksamkeitstest – München*. Göttingen: Hogrefe.
- Schall, S. (2004). Forensische Linguistik. In K. Knapp, G. Antos, M. Becker-Mrotzek, A. Deppermann, S. Göpferich, J. Grabowski, M. Klemm & C. Villiger (Hrsg.). *Angewandte Linguistik*. Tübingen, Basel: A. Francke Verlag, pp. 544-562.
- Scherer, K.R. (1974). Voice quality analysis of American and German speakers. *Journal of Psycholinguistic Research*, 3:281-298.
- Schneider, B., Enne, R., Cecon, M., Diendorfer-Radner, G., Wittels, P., Bigenzahn, W. & Johannes, B. (2006). Effects of vocal constitution and autonomic stress-related reactivity on vocal endurance in female student teachers. *Journal of Voice*, 20:242-250.
- Schneider, B., van Trotsenburg, M., Hanke, G., Bigenzahn, W. & Huber, J. (2004). Voice impairment and menopause. *Menopause*, 11:151-158.
- Schnider, A., Benson, D.F., Alexander, D.N. & Schnider-Klaus, A. (1994). Non-verbal environmental sound recognition after unilateral hemispheric stroke. *Brain: A Journal of Neurology*, 117:281-287.
- Schöler, M. & Grötzbach, H. (2002). *Aphasie – Wege aus dem Sprachdschungel*. Berlin, Heidelberg, New York, Barcelona, Hongkong, London, Mailand, Paris, Tokio: Springer.

- Schweinberger, S.R., Herholz, A. & Sommer W. (1997). Recognizing famous voices: influence of stimulus duration and different types of retrieval cues. *Journal of Speech, Language, and Hearing Research*, 40:453-463.
- Scott, S.K., Young, A.W., Calder, A.J., Hellawell, D.J., Aggleton, J.P. & Johnson, M. (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature*, 385:254-257.
- Shah, N.J., Marshall, J.C., Zafiris, O., Schwab, A., Zilles, K., Markowitsch, H.J. & Fink, G.R. (2001). The neural correlates of person familiarity. A functional magnetic resonance imaging study with clinical implications. *Brain: A Journal of Neurology*, 124:804-815.
- Shipp, T. & Hollien, H. (1969). Perception of the aging male voice. *Journal of Speech and Hearing Research*, 12:703-710.
- Small, D.M., Bernasconi, N., Bernasconi, A., Sziklas, V. & Jones-Gotman, M. (2005). Gustatory agnosia. *Neurology*, 64:311-317.
- Sodersten, M., Granqvist, S., Hammarberg, B. & Szabo, A. (2002). Vocal behavior and vocal loading factors for preschool teachers at work studied with binaural DAT recordings. *Journal of Voice*, 16:356-371.
- Spence, M.J., Rollins, P.R. & Jerger S. (2002). Children's recognition of cartoon voices. *Journal of Speech, Language, and Hearing Research*, 45:214-222.
- Spencer, L.E. (1988). Speech characteristics of male-to-female transsexuals: a perceptual and acoustic study. *Folia Phoniatrica*, 40:31-42.
- Spreen, O., Benton, A.L. & Fincham, R.W. (1965). Auditory Agnosia without Aphasia. *Archives of Neurology*, 13:84-92.
- Steeves, J.K., Culham, J.C., Duchaine, B.C., Pratesi, C.C., Valyear, K.F., Schindler, I., Humphrey, G.K., Milner, A.D. & Goodale, M.A. (2006). The fusiform face area is not sufficient for face recognition: evidence from a patient with dense prosopagnosia and no occipital face area. *Neuropsychologia*, 44:594-609.
- Steller, S., Wellershaus, P. & Wolf, T. (1992). Realkennzeichen in Kinderaussagen: Empirische Grundlagen der Kriterienorientierten Aussageanalyse. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 39:151-170.

- Szirmai, I., Farsang, M. & Csuri, M. (2003). Cortical auditory disorder caused by bilateral strategic cerebral bleedings. Analyses of two cases. *Brain and Language*, 85:159-165.
- Taniwaki, T., Tagawa, K., Sato, F. & Iino, K. (2000). Auditory agnosia restricted to environmental sounds following cortical deafness and generalized auditory agnosia. *Clinical Neurology and Neurosurgery*, 102:156-162.
- Tesak, J. (2006). *Einführung in die Aphasologie*. Stuttgart, New York: Thieme.
- Thompson, C.P. (1987). A language effect in voice recognition. *Applied Cognitive Psychology*, 1:121-131.
- Titze, I.R. (1989). On the relation between subglottal pressure and fundamental frequency in phonation. *The Journal of the Acoustical Society of America*, 85:901-906.
- Titze, I.R. (1991). A model for neurologic sources of aperiodicity in vocal fold vibration. *Journal of Speech and Hearing Research*, 34:460-472.
- Titze, I.R. (1994). *Principles of voice production*. London, Sydney, Toronto, Mexico, New Delhi, Tokyo, Singapore, Rio de Janeiro: Prentice Hall.
- Titze, I.R. (1995). Regulation of fundamental frequency with a physiologically-based model of the larynx. In S. Kiritani, H. Hirose & H. Fujisaki (Eds.). *Speech Production and Language – In Honor of Osamu Fujimura*. Berlin, New York: Mouton de Gruyter. pp. 33-51.
- Tosi, O. (1979). *Voice identification: Theory and Legal Applications*. Baltimore: University Park Press.
- Tosi, O., Oyer, H., Lashbrook, W., Pedrey, C., Nicol, J. & Nash, E. (1972). Experiment on voice identification. *Journal of the Acoustical Society of America*, 51:2030-2043.
- Tramo, M.J., Shah, G.D. & Braida, L.D. (2002). Functional role of auditory cortex in frequency processing and pitch perception. *Journal of Neurophysiology*, 87:122-139.
- Turnure, C. (1971). Response to voice of mother and stranger by babies in the first year. *Developmental Psychology*, 4:182-190.

- Uttner, I., Mottaghy, F.M., Schreiber, H., Riecker, A., Ludolph, A.C. & Kas-subek, J. (2006). Primary progressive aphasia accompanied by environmental sound agnosia: A neuropsychological, MRI and PET study. *Psychiatry research*, 146:191-197.
- Valentine, T., Brédart, S., Lawson, R. & Ward, G. (1991). What's in a name? Access to information from people's names. *European Journal of Cognitive Psychology*, 3:147-176.
- Valentine, T. & Bruce, V. (1986). The effects of distinctiveness in recognising and classifying faces. *Perception*, 15:525-535.
- Van Dommelen, W.A. (1990). Acoustic parameters in human speaker recognition. *Language and Speech*, 33:259-272.
- Van Lancker, D.R. & Canter, G.J. (1982). Impairment of voice and face recognition in patients with hemispheric damage. *Brain and Cognition*, 1:185-195.
- Van Lancker, D.R., Cummings, J.L., Kreiman, J. & Dobkin, B.H. (1988). Phonagnosia: a dissociation between familiar and unfamiliar voices. *Cortex*, 24:195-209.
- Van Lancker, D.R. & Kreiman, J. (1987). Voice discrimination and recognition are separate abilities. *Neuropsychologia*, 25:829-834.
- Van Lancker, D.R., Kreiman, J. & Cummings, J. (1989). Voice perception deficits: neuroanatomical correlates of phonagnosia. *Journal of Clinical and Experimental Neuropsychology*, 11:665-674.
- Van Lancker, D.R., Kreiman, J. & Emmorey, K. (1985). Familiar voice recognition: patterns and parameters – Part I: Recognition of backward voices. *Journal of Phonetics*, 13:19-38.
- Van Lancker, D.R., Kreiman, J. & Wickens, T. (1985). Familiar voice recognition: patterns and parameters – Part II: Recognition of rate-altered voices. *Journal of Phonetics*, 13:39-52.
- Van Lancker, D.R. & Sidtis, J.J. (1992). The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: all errors are not created equal. *Journal of Speech and Hearing Research*, 35:963-970.
- Verdonck-de Leeuw, I.M. & Mahieu, H.F. (2004). Vocal aging and the impact on daily life: a longitudinal study. *Journal of Voice*, 18:193-202.

- Vignolo, L.A. (1969). Auditory agnosia: a review and report of recent evidence. In A.L. Benton (Ed.), *Contributions to clinical neuropsychology*. Chicago: Aldine, pp. 172-208.
- Vignolo LA. (2003). Music agnosia and auditory agnosia. Dissociations in stroke patients. *Annals of the New York Academy of Sciences*, 999:50-57.
- Voiers, W.D. (1964). Perceptual bases of speaker identity. *The Journal of the Acoustical Society of America*, 36:1065-1073.
- von Kriegstein, K., Eger, E., Kleinschmidt, A. & Giraud, A.L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cognitive Brain Research*, 17:48-55.
- von Kriegstein, K. & Giraud, A.L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *NeuroImage*, 22:948-955.
- Walden, B.R., Montgomery, A.A., Gibeily, G.J., Prosek, R.A. & Schwartz, D.M. (1978). Correlates of psychological dimensions in talker similarity. *Journal of Speech and Hearing Research*, 21:265-275.
- Warnick, D. & Sanders, G. (1980). Why do eyewitnesses make so many mistakes? *Journal of Applied Social Psychology*, 10:362-366.
- Warren, C. & Morton, J. (1982). The effects of priming on picture recognition. *British Journal of Psychology*, 73:117-129.
- Warrier, C.M. & Zatorre, R.J. (2004). Right temporal cortex is critical for utilization of melodic contextual cues in a pitch constancy task. *Brain : A Journal of Neurology*, 127:1616-1625.
- Wechsler, D. (1987). *WMS-R: Wechsler Memory Scale - Revised (Manual)*. San Antonio: The Psychological Corporation.
- Weinberg, B. & Bennett, S. (1971). Speaker sex recognition of 5- and 6-year-old children's voices. *The Journal of the Acoustical Society of America*, 50:1210-1213.
- Whiteside, S.P. (1998). Identification of a speaker's sex: a study of vowels. *Perceptual and Motor Skills*, 86:579-584.
- Whiteside, S.P. (1999). Acoustic characteristics of vocal emotions simulated by actors. *Perceptual and Motor Skills*, 89:1195-1208.

- Wildgruber, D., Pihan, H., Ackermann, H., Erb, M. & Grodd, W. (2002). Dynamic brain activation during processing of emotional intonation: influence of acoustic parameters, emotional valence, and sex. *NeuroImage*, 15:856-869.
- Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T. & Ackermann H. (2005). Identification of emotional intonation evaluated by fMRI. *NeuroImage*, 24:1233-1241.
- Wilding, J. & Cook, S. (2000). Sex differences and individual consistency in voice identification. *Perceptual and Motor Skills*, 91:535-538.
- Willi, U. (⁴2001). Phonetik und Phonologie. In A. Linke, M. Nussbaumer & P.R. Portmann. *Studienbuch Linguistik*. Tübingen: Niemeyer, pp. 401-435.
- Yarmey, A.D. (1991). Voice identification over the telephone. *Journal of Applied Social Psychology*, 22:1868-1876.
- Yarmey, A.D. & Matthys, E. (1992). Voice identification of an abductor. *Applied Cognitive Psychology*, 6:367-377.
- Zatorre, R.J., Evans, A.C., Meyer, E. & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, 256:846-849.
- Zigmond, A.S. & Snaith, R.P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67:361-370.
- Zwicker, E. & Feldtkeller, R. (²1967). *Das Ohr als Nachrichtenempfänger*. Stuttgart, Leipzig: S. Hirzel.

Teil IV.

Anhang

12. Tabellen und Abbildungen

12.1. Tabellen

| PbNr. | Kürzel | Geschl. | Alter | p.o.-time | Syndrom |
|-------|--------|---------|-------|-----------|----------|
| 1 | LBD 01 | m | 64,1 | 1 | Wernicke |
| 2 | LBD 02 | m | 60,4 | 2 | Wernicke |
| 3 | LBD 03 | m | 73,9 | 7 | Amnest. |
| 4 | LBD 04 | m | 61,5 | 1 | Broca |
| 5 | LBD 05 | m | 72,6 | 2 | Amnest. |
| 6 | LBD 06 | m | 64,7 | 1 | Amnest. |
| 7 | LBD 07 | m | 50,1 | 1 | Amnest. |
| 8 | LBD 08 | w | 61,8 | 12 | Amnest. |
| 9 | LBD 09 | m | 71,6 | 1 | Wernicke |
| 10 | LBD 10 | w | 73,3 | 1 | Wernicke |
| 11 | LBD 11 | m | 74,9 | 5 | Wernicke |

Tabelle 12.1.: Deskriptive Statistik der Gruppe 1

| PbNr. | Kürzel | Geschl. | Alter | p.o.-time |
|-------|--------|---------|-------|-----------|
| 12 | RBD 01 | w | 58,3 | 4 |
| 13 | RBD 02 | m | 68,4 | 1 |
| 14 | RBD 03 | m | 50,1 | 1 |
| 15 | RBD 04 | m | 67,9 | 2 |
| 16 | RBD 05 | m | 59,0 | 1 |
| 17 | RBD 06 | w | 77,8 | 2 |
| 18 | RBD 07 | w | 49,6 | 1 |
| 19 | RBD 08 | m | 64,5 | 1 |
| 20 | RBD 09 | m | 81,4 | 1 |

Tabelle 12.2.: Deskriptive Statistik der Gruppe 2

12. Tabellen und Abbildungen

| PbNr. | Kürzel | Geschl. | Alter |
|-------|--------|---------|-------|
| 21 | HC 01 | m | 66,7 |
| 22 | HC 02 | m | 63,3 |
| 23 | HC 03 | m | 72,5 |
| 24 | HC 04 | m | 63,3 |
| 25 | HC 05 | m | 66,0 |
| 26 | HC 06 | w | 63,1 |
| 27 | HC 07 | w | 61,5 |
| 28 | HC 08 | w | 64,1 |
| 29 | HC 09 | w | 61,6 |
| 30 | HC 10 | w | 65,4 |
| 31 | HC 11 | w | 65,7 |
| 32 | HC 12 | m | 69,9 |
| 33 | HC 13 | w | 67,6 |
| 34 | HC 14 | w | 66,9 |
| 35 | HC 15 | m | 67,0 |
| 36 | HC 16 | w | 51,7 |
| 37 | HC 17 | m | 53,1 |

Tabelle 12.3.: Deskriptive Statistik der Gruppe 3

12. Tabellen und Abbildungen

12. Tabellen und Abbildungen

| PbNr. | Kürzel | WMS-R | AKT-T | AKT-R | AKT-F% | AKT-G | HADS-D |
|-------|--------|-------|-------|-------|--------|-------|--------|
| 1 | LBD 01 | 15 | 66,9 | 78,3 | 100,0 | 81,5 | 4 |
| 2 | LBD 02 | 19 | 77,7 | 100,0 | 100,0 | 100,0 | 2 |
| 3 | LBD 03 | 21 | 77,7 | 100,0 | 100,0 | 100,0 | 1 |
| 4 | LBD 04 | 16 | 80,4 | 100,0 | 100,0 | 100,0 | 6 |
| 5 | LBD 05 | 21 | 80,4 | 78,3 | 100,0 | 81,5 | 1 |
| 6 | LBD 06 | 21 | 98,4 | 100,0 | 100,0 | 100,0 | 1 |
| 7 | LBD 07 | 19 | 98,4 | 100,0 | 100,0 | 100,0 | 1 |
| 8 | LBD 08 | 18 | 93,5 | 100,0 | 100,0 | 100,0 | 8 |
| 9 | LBD 09 | 19 | 96,7 | 100,0 | 100,0 | 100,0 | 5 |
| 10 | LBD 10 | 18 | 96,2 | 78,3 | 100,0 | 81,5 | 1 |
| 11 | LBD 11 | 22 | 96,7 | 60,9 | 100,0 | 67,9 | 6 |
| 12 | RBD 01 | 21 | 97,4 | 48,6 | 100,0 | 80,4 | 2 |
| 13 | RBD 02 | 21 | 80,8 | 85,1 | 100,0 | 96,1 | 5 |
| 14 | RBD 03 | 22 | 99,0 | 75,5 | 100,0 | 92,4 | 5 |
| 15 | RBD 04 | 20 | 94,0 | 55,1 | 78,9 | 80,4 | 1 |
| 16 | RBD 05 | 18 | 73,2 | 63,7 | 100,0 | 87,7 | 4 |
| 17 | RBD 06 | 24 | 73,2 | 63,7 | 30,8 | 39,7 | 1 |
| 18 | RBD 07 | 16 | 87,4 | 85,1 | 100,0 | 96,1 | 2 |
| 19 | RBD 08 | 17 | 80,8 | 100,0 | 100,0 | 100,0 | 9 |
| 20 | RBD 09 | 21 | 21,1 | 85,1 | 72,6 | 84,0 | 1 |
| 21 | HC 01 | 20 | 98,9 | 100,0 | 100,0 | 100,0 | 0 |
| 22 | HC 02 | 18 | 96,4 | 54,7 | 100,0 | 69,4 | 4 |
| 23 | HC 03 | 18 | 92,5 | 75,9 | 100,0 | 83,9 | 2 |
| 24 | HC 04 | 18 | 96,4 | 75,9 | 100,0 | 83,9 | 1 |
| 25 | HC 05 | 19 | 69,2 | 75,9 | 100,0 | 83,9 | 3 |
| 26 | HC 06 | 18 | 100,0 | 100,0 | 100,0 | 100,0 | 1 |
| 27 | HC 07 | 18 | 92,5 | 75,9 | 100,0 | 83,9 | 0 |
| 28 | HC 08 | 19 | 96,4 | 100,0 | 100,0 | 100,0 | 2 |
| 29 | HC 09 | 22 | 96,4 | 75,9 | 100,0 | 83,9 | 1 |
| 30 | HC 10 | 23 | 92,5 | 100,0 | 100,0 | 100,0 | 3 |
| 31 | HC 11 | 19 | 69,2 | 75,9 | 49,2 | 69,4 | 6 |
| 32 | HC 12 | 21 | 96,4 | 100,0 | 100,0 | 100,0 | 6 |
| 33 | HC 13 | 20 | 100,0 | 100,0 | 100,0 | 100,0 | 1 |
| 34 | HC 14 | 20 | 82,1 | 100,0 | 100,0 | 100,0 | 2 |
| 35 | HC 15 | 23 | 96,4 | 100,0 | 100,0 | 100,0 | 6 |
| 36 | HC 16 | 21 | 82,1 | 100,0 | 100,0 | 100,0 | 0 |
| 37 | HC 17 | 19 | 58,7 | 100,0 | 100,0 | 100,0 | 0 |

Tabelle 12.4.: Erreichte Rohwerte (HADS-D / WMS-R) bzw. Prozentränge (AKT)

12. Tabellen und Abbildungen

| Distraktor | Median Pitch | Mean Pitch | SD | min. Pitch | max. Pitch | Jitter (%) | Shimmer (%) |
|------------|--------------|------------|-------|------------|------------|------------|-------------|
| wD1 | 218,46 | 225,88 | 64,38 | 76,23 | 485,98 | 2,91 | 10,56 |
| wD2 | 233,55 | 233,25 | 44,28 | 83,87 | 408,27 | 2,55 | 10,15 |
| wD3 | 177,16 | 186,03 | 54,88 | 75,77 | 487,57 | 2,73 | 12,28 |
| wD4 | 218,72 | 225,90 | 48,26 | 105,52 | 353,65 | 1,96 | 9,38 |
| wD5 | 179,75 | 175,46 | 25,81 | 98,45 | 244,41 | 2,25 | 13,79 |
| mD1 | 118,15 | 119,64 | 25,05 | 75,46 | 202,40 | 2,74 | 14,45 |
| mD2 | 148,84 | 150,91 | 36,12 | 84,94 | 394,70 | 2,95 | 12,36 |
| mD3 | 180,12 | 175,56 | 26,94 | 89,57 | 244,41 | 2,41 | 13,43 |
| mD4 | 120,27 | 126,22 | 21,33 | 86,79 | 198,20 | 3,50 | 12,04 |
| mD5 | 177,16 | 186,03 | 54,88 | 75,74 | 487,57 | 2,73 | 12,28 |

Tabelle 12.5.: Akustische Parameter der Distraktorstimmen

12. Tabellen und Abbildungen

| Item | min. | max. | Mittelwert | SD | | Erk. | SD |
|------|------|------|------------|------|--|------|------|
| V1 | -2 | 1 | -,09 | ,831 | | ,73 | ,467 |
| V2 | -2 | 1 | -,64 | ,924 | | ,55 | ,522 |
| CVC1 | -1 | 1 | ,18 | ,751 | | ,73 | ,467 |
| CVC2 | -2 | 1 | -,64 | ,924 | | ,91 | ,302 |
| W1 | 0 | 1 | ,45 | ,522 | | ,91 | ,302 |
| W2 | 0 | 2 | 1,27 | ,786 | | ,91 | ,302 |
| S1 | 1 | 2 | 1,82 | ,405 | | 1,0 | ,000 |
| S2 | 1 | 2 | 1,82 | ,405 | | 1,0 | ,000 |

Tabelle 12.6.: Schwierigkeitsgrad vs. Stimmerkennung (LBD)

| Item | min. | max. | Mittelwert | SD | | Erk. | SD |
|------|------|------|------------|------|--|------|------|
| V1 | 0 | 2 | ,82 | ,751 | | ,73 | ,467 |
| V2 | -1 | 1 | ,09 | ,831 | | ,55 | ,522 |
| CVC1 | 0 | 2 | ,91 | ,539 | | ,91 | ,302 |
| CVC2 | -1 | 2 | ,18 | ,874 | | ,73 | ,467 |
| W1 | 0 | 2 | 1,0 | ,775 | | ,91 | ,302 |
| W2 | 0 | 2 | 1,36 | ,674 | | ,91 | ,302 |
| S1 | 1 | 2 | 1,91 | ,302 | | 1,0 | ,000 |
| S2 | 1 | 2 | 1,91 | ,302 | | 1,0 | ,000 |

Tabelle 12.7.: Sicherheitsgefühl vs. Stimmerkennung (LBD)

12. Tabellen und Abbildungen

| Item | min. | max. | Mittelwert | SD | | Erk. | SD |
|------|------|------|------------|-------|--|------|------|
| V1 | -2 | 0 | -1,11 | ,928 | | ,44 | ,527 |
| V2 | -2 | 1 | -1,0 | 1,118 | | ,33 | ,500 |
| CVC1 | -2 | 2 | -,33 | 1,323 | | ,78 | ,441 |
| CVC2 | -2 | 1 | -,44 | 1,130 | | ,44 | ,527 |
| W1 | -2 | 2 | ,22 | 1,202 | | ,33 | ,500 |
| W2 | -1 | 2 | ,56 | 1,014 | | ,56 | ,527 |
| S1 | -1 | 2 | ,67 | 1,0 | | 1,0 | ,000 |
| S2 | -1 | 2 | 1,0 | 1,118 | | 1,0 | ,000 |

Tabelle 12.8.: Schwierigkeitsgrad vs. Stimmerkennung (RBD)

| Item | min. | max. | Mittelwert | SD | | Erk. | SD |
|------|------|------|------------|-------|--|------|------|
| V1 | -2 | 2 | -,33 | 1,658 | | ,44 | ,527 |
| V2 | -2 | 2 | -,33 | 1,225 | | ,33 | ,500 |
| CVC1 | -1 | 2 | ,56 | ,882 | | ,78 | ,441 |
| CVC2 | -1 | 2 | ,44 | ,882 | | ,44 | ,527 |
| W1 | 0 | 2 | ,89 | ,782 | | ,33 | ,500 |
| W2 | 0 | 2 | ,89 | ,782 | | ,56 | ,527 |
| S1 | 0 | 2 | 1,44 | ,882 | | 1,0 | ,000 |
| S2 | 0 | 2 | 1,67 | ,707 | | 1,0 | ,000 |

Tabelle 12.9.: Sicherheitsgefühl vs. Stimmerkennung (RBD)

12. Tabellen und Abbildungen

| Item | min. | max. | Mittelwert | SD | sE + vE | SD |
|------|------|------|------------|------|---------|------|
| V1 | -2 | 1 | -,47 | ,943 | ,76 | ,437 |
| V2 | -2 | 1 | -,71 | ,985 | ,53 | ,514 |
| CVC1 | -1 | 1 | ,06 | ,748 | ,65 | ,493 |
| CVC2 | -2 | 1 | -,65 | ,862 | ,53 | ,514 |
| W1 | -1 | 1 | -,06 | ,827 | ,65 | ,493 |
| W2 | -1 | 2 | ,65 | ,786 | ,94 | ,243 |
| S1 | 0 | 2 | 1,24 | ,664 | 1,0 | ,000 |
| S2 | 1 | 2 | 1,47 | ,624 | 1,0 | ,000 |

Tabelle 12.10.: Schwierigkeitsgrad vs. Stimmerkennung (HC)

| Item | min. | max. | Mittelwert | SD | sE + vE | SD |
|------|------|------|------------|-------|---------|------|
| V1 | -2 | 2 | ,35 | 1,169 | ,68 | ,475 |
| V2 | -2 | 2 | ,06 | 1,144 | ,53 | ,514 |
| CVC1 | -1 | 2 | ,71 | 1,105 | ,65 | ,493 |
| CVC2 | -1 | 2 | ,00 | ,866 | ,53 | ,514 |
| W1 | -1 | 2 | ,35 | 1,057 | ,65 | ,493 |
| W2 | -1 | 2 | 1,06 | ,966 | ,94 | ,243 |
| S1 | 0 | 2 | 1,71 | ,588 | 1,0 | ,000 |
| S2 | 1 | 2 | 1,76 | ,437 | 1,0 | ,000 |

Tabelle 12.11.: Sicherheitsgefühl vs. Stimmerkennung (HC)

12. Tabellen und Abbildungen

| Item | min. | max. | Mittelwert | SD | sE + vE | SD |
|------|------|------|------------|------|---------|------|
| V1 | -2 | 1 | -,51 | ,961 | ,68 | ,475 |
| V2 | -2 | 1 | -,76 | ,983 | ,49 | ,507 |
| CVC1 | -2 | 2 | ,00 | ,913 | ,76 | ,435 |
| CVC2 | -2 | 1 | -,59 | ,927 | ,57 | ,502 |
| W1 | -2 | 2 | ,16 | ,866 | ,65 | ,484 |
| W2 | -1 | 2 | ,81 | ,877 | ,84 | ,374 |
| S1 | -1 | 2 | 1,27 | ,804 | 1,0 | ,000 |
| S2 | -1 | 2 | 1,46 | ,767 | 1,0 | ,000 |

Tabelle 12.12.: Schwierigkeitsgrad vs. Stimmerkennung (Gesamt)

| Item | min. | max. | Mittelwert | SD | sE + vE | SD |
|------|------|------|------------|-------|---------|------|
| V1 | -2 | 2 | ,32 | 1,248 | ,68 | ,475 |
| V2 | -2 | 2 | -,03 | 1,067 | ,49 | ,507 |
| CVC1 | -1 | 2 | ,73 | ,902 | ,76 | ,435 |
| CVC2 | -1 | 2 | ,16 | ,866 | ,57 | ,502 |
| W1 | -1 | 2 | ,68 | ,944 | ,65 | ,484 |
| W2 | -1 | 2 | 1,11 | ,843 | ,84 | ,374 |
| S1 | 0 | 2 | 1,70 | ,618 | 1,0 | ,000 |
| S2 | 0 | 2 | 1,78 | ,479 | 1,0 | ,000 |

Tabelle 12.13.: Sicherheitsgefühl vs. Stimmerkennung (Gesamt)

12. Tabellen und Abbildungen

12.2. Abbildungen

| | | Quadratsumme | df | Mittel der Quadrate | F | Signifikanz |
|-----------|-----------------------|--------------|----|---------------------|-------|-------------|
| Alter | Zwischen den Gruppen | 36,844 | 2 | 18,422 | ,311 | ,735 |
| | Innerhalb der Gruppen | 2016,353 | 34 | 59,305 | | |
| | Gesamt | 2053,197 | 36 | | | |
| Monate_po | Zwischen den Gruppen | 11,669 | 1 | 11,669 | 1,554 | ,228 |
| | Innerhalb der Gruppen | 135,131 | 18 | 7,507 | | |
| | Gesamt | 146,800 | 19 | | | |

Abbildung 12.1.: ANOVA zu den Variablen Alter und post-onset-time

| | | Quadratsumme | df | Mittel der Quadrate | F | Signifikanz |
|----------------|-----------------------|--------------|----|---------------------|------|-------------|
| HADS_D_Rohwert | Zwischen den Gruppen | 10,435 | 2 | 5,218 | ,890 | ,420 |
| | Innerhalb der Gruppen | 199,241 | 34 | 5,860 | | |
| | Gesamt | 209,676 | 36 | | | |
| WMS_R_Rohwert | Zwischen den Gruppen | 5,860 | 2 | 2,930 | ,677 | ,515 |
| | Innerhalb der Gruppen | 147,059 | 34 | 4,325 | | |
| | Gesamt | 152,919 | 36 | | | |

Abbildung 12.2.: ANOVA zu den Variablen HADS-D und WMS-R

| | | Quadratsumme | df | Mittel der Quadrate | F | Signifikanz |
|----------|-----------------------|--------------|----|---------------------|-------|-------------|
| AKT_T_PR | Zwischen den Gruppen | 698,273 | 2 | 349,137 | 1,448 | ,249 |
| | Innerhalb der Gruppen | 8196,394 | 34 | 241,070 | | |
| | Gesamt | 8894,668 | 36 | | | |
| AKT_R_PR | Zwischen den Gruppen | 1752,371 | 2 | 876,185 | 3,904 | ,030 |
| | Innerhalb der Gruppen | 7630,639 | 34 | 224,431 | | |
| | Gesamt | 9383,010 | 36 | | | |
| AKT_F_PR | Zwischen den Gruppen | 923,699 | 2 | 461,849 | 2,284 | ,117 |
| | Innerhalb der Gruppen | 6874,193 | 34 | 202,182 | | |
| | Gesamt | 7797,892 | 36 | | | |
| AKT_G_PR | Zwischen den Gruppen | 407,028 | 2 | 203,514 | 1,159 | ,326 |
| | Innerhalb der Gruppen | 5968,613 | 34 | 175,547 | | |
| | Gesamt | 6375,641 | 36 | | | |

Abbildung 12.3.: ANOVA zu den Variablen des AKT

12. Tabellen und Abbildungen

PC_passed

| | Quadratsumme | df | Mittel der Quadrate | F | Signifikanz |
|-----------------------|--------------|----|---------------------|-------|-------------|
| Zwischen den Gruppen | 2650,170 | 2 | 1325,085 | 3,516 | ,041 |
| Innerhalb der Gruppen | 12814,357 | 34 | 376,893 | | |
| Gesamt | 15464,527 | 36 | | | |

Abbildung 12.4.: ANOVA zur Variable *percentage correct (PC)*

| | Gruppe | N | Mittelwert | Standardabweichung | Standardfehler des Mittelwertes |
|-----------|--------|----|------------|--------------------|---------------------------------|
| PC_passed | LBD | 11 | 84,091 | 15,9010 | 4,7943 |
| | RBD | 9 | 61,111 | 20,1987 | 6,7329 |
| sE_ges | LBD | 11 | 77,273 | 17,5162 | 5,2813 |
| | RBD | 9 | 54,167 | 18,7500 | 6,2500 |
| vE_ges | LBD | 11 | 6,818 | 11,6775 | 3,5209 |
| | RBD | 9 | 6,944 | 9,0810 | 3,0270 |

Abbildung 12.5.: Deskriptive Statistik der Mittelwerte der Gruppen LBD und RBD

12. Tabellen und Abbildungen

| | sE_ges | vE_ges | nE_ges |
|-------------------------------------------|-------------------|-------------------|-------------------|
| Mann-Whitney-U | 18,500 | 46,000 | 19,000 |
| Wilcoxon-W | 63,500 | 112,000 | 85,000 |
| Z | -2,407 | -,305 | -2,367 |
| Asymptotische Signifikanz (2-seitig) | ,016 | ,760 | ,018 |
| Exakte Signifikanz [2*(1-seitig Sig.)] | ,016 ^a | ,824 ^a | ,020 ^a |

a. Nicht für Bindungen korrigiert.

Abbildung 12.6.: U-Test: Erkennungsleistung der Gruppen LBD und RBD

| | Gruppe | N | Mittelwert | Standardabweichung | Standardfehler des Mittelwertes |
|-----------|--------|----|------------|--------------------|---------------------------------|
| PC_passed | RBD | 9 | 61,111 | 20,1987 | 6,7329 |
| | HC | 17 | 75,735 | 20,9494 | 5,0810 |
| sE_ges | RBD | 9 | 54,167 | 18,7500 | 6,2500 |
| | HC | 17 | 65,441 | 20,9768 | 5,0876 |
| vE_ges | RBD | 9 | 6,944 | 9,0810 | 3,0270 |
| | HC | 17 | 10,294 | 7,9492 | 1,9280 |

Abbildung 12.7.: Deskriptive Statistik der Mittelwerte der Gruppen RBD und HC

12. Tabellen und Abbildungen

| | sE_ges | vE_ges | nE_ges |
|----------------------------------------|-------------------|-------------------|-------------------|
| Mann-Whitney-U | 51,500 | 58,500 | 46,000 |
| Wilcoxon-W | 96,500 | 103,500 | 199,000 |
| Z | -1,370 | -1,073 | -1,677 |
| Asymptotische Signifikanz (2-seitig) | ,171 | ,283 | ,093 |
| Exakte Signifikanz [2*(1-seitig Sig.)] | ,181 ^a | ,339 ^a | ,107 ^a |

a. Nicht für Bindungen korrigiert.

Abbildung 12.8.: U-Test: Erkennungsleistung der Gruppen RBD und HC

| | Gruppe | N | Mittelwert | Standardabweichung | Standardfehler des Mittelwertes |
|-----------|--------|----|------------|--------------------|---------------------------------|
| PC_passed | LBD | 11 | 84,091 | 15,9010 | 4,7943 |
| | HC | 17 | 75,735 | 20,9494 | 5,0810 |
| sE_ges | LBD | 11 | 77,273 | 17,5162 | 5,2813 |
| | HC | 17 | 65,441 | 20,9768 | 5,0876 |
| vE_ges | LBD | 11 | 6,818 | 11,6775 | 3,5209 |
| | HC | 17 | 10,294 | 7,9492 | 1,9280 |

Abbildung 12.9.: Deskriptive Statistik der Mittelwerte der Gruppen LBD und HC

| | sE_ges | vE_ges | nE_ges |
|----------------------------------------|-------------------|-------------------|-------------------|
| Mann-Whitney-U | 62,000 | 64,500 | 72,500 |
| Wilcoxon-W | 215,000 | 130,500 | 138,500 |
| Z | -1,510 | -1,505 | -1,020 |
| Asymptotische Signifikanz (2-seitig) | ,131 | ,132 | ,308 |
| Exakte Signifikanz [2*(1-seitig Sig.)] | ,147 ^a | ,175 ^a | ,329 ^a |

a. Nicht für Bindungen korrigiert.

Abbildung 12.10.: U-Test: Erkennungsleistung der Gruppen LBD und HC

12. Tabellen und Abbildungen

PC passed

| | Quadratsumme | df | Mittel der Quadrate | F | Signifikanz |
|-----------------------|--------------|----|---------------------|------|-------------|
| Zwischen den Gruppen | 153,409 | 2 | 76,705 | ,258 | ,779 |
| Innerhalb der Gruppen | 2375,000 | 8 | 296,875 | | |
| Gesamt | 2528,409 | 10 | | | |

Abbildung 12.11.: Mittelwertvergleich zwischen den Leistungen der Amnestischen und der Wernicke-Aphasiker

| | Gruppe | Mittelwert | Standardabweichung | N |
|--------------|--------|------------|--------------------|----|
| V_ges_Proz | LBD | 63,64 | 39,312 | 11 |
| | RBD | 38,89 | 41,667 | 9 |
| | HC | 64,71 | 42,444 | 17 |
| | Gesamt | 58,11 | 41,689 | 37 |
| CVC_ges_Proz | LBD | 81,82 | 25,226 | 11 |
| | RBD | 61,11 | 33,333 | 9 |
| | HC | 58,82 | 36,380 | 17 |
| | Gesamt | 66,22 | 33,446 | 37 |
| W_ges_Proz | LBD | 90,91 | 20,226 | 11 |
| | RBD | 44,44 | 39,087 | 9 |
| | HC | 79,41 | 25,365 | 17 |
| | Gesamt | 74,32 | 32,535 | 37 |
| S_ges_Proz | LBD | 100,00 | ,000 | 11 |
| | RBD | 100,00 | ,000 | 9 |
| | HC | 100,00 | ,000 | 17 |
| | Gesamt | 100,00 | ,000 | 37 |

Abbildung 12.12.: Deskriptive Statistik der Erkennungsleistung (*percentage correct [PC]*)

12. Tabellen und Abbildungen

Maß: MASS_1

| Quelle | | Quadratsumme vom Typ III | df | Mittel der Quadrate | F | Signifikanz | Partielles Eta-Quadrat |
|---------------|------------------------|--------------------------|---------|---------------------|--------|-------------|------------------------|
| styp | Sphärizität angenommen | 36578,991 | 3 | 12192,997 | 17,781 | ,000 | ,343 |
| | Greenhouse-Geisser | 36578,991 | 2,565 | 14263,101 | 17,781 | ,000 | ,343 |
| | Huynh-Feldt | 36578,991 | 2,955 | 12378,924 | 17,781 | ,000 | ,343 |
| | Untergrenze | 36578,991 | 1,000 | 36578,991 | 17,781 | ,000 | ,343 |
| styp * Gruppe | Sphärizität angenommen | 9142,056 | 6 | 1523,676 | 2,222 | ,047 | ,116 |
| | Greenhouse-Geisser | 9142,056 | 5,129 | 1782,363 | 2,222 | ,058 | ,116 |
| | Huynh-Feldt | 9142,056 | 5,910 | 1546,910 | 2,222 | ,048 | ,116 |
| | Untergrenze | 9142,056 | 2,000 | 4571,028 | 2,222 | ,124 | ,116 |
| Fehler(styp) | Sphärizität angenommen | 69945,781 | 102 | 685,743 | | | |
| | Greenhouse-Geisser | 69945,781 | 87,196 | 802,167 | | | |
| | Huynh-Feldt | 69945,781 | 100,468 | 696,200 | | | |
| | Untergrenze | 69945,781 | 34,000 | 2057,229 | | | |

Abbildung 12.13.: Tests der Innersubjekteffekte

Maß: MASS_1

Transformierte Variable: Mittel

| Quelle | Quadratsumme vom Typ III | df | Mittel der Quadrate | F | Signifikanz | Partielles Eta-Quadrat |
|-----------------|--------------------------|----|---------------------|---------|-------------|------------------------|
| Konstanter Term | 748544,680 | 1 | 748544,68 | 496,524 | ,000 | ,936 |
| Gruppe | 10600,681 | 2 | 5300,340 | 3,516 | ,041 | ,171 |
| Fehler | 51257,427 | 34 | 1507,571 | | | |

Abbildung 12.14.: Tests der Zwischensubjekteffekte

12. Tabellen und Abbildungen

| | CVC_ges_ Proz |
|-------------------------------------------|-------------------|
| Mann-Whitney-U | 61,000 |
| Wilcoxon-W | 214,000 |
| Z | -1,687 |
| Asymptotische Signifikanz (2-seitig) | ,092 |
| Exakte Signifikanz [2*(1-seitig Sig.)] | ,134 ^a |

a. Nicht für Bindungen korrigiert.

Abbildung 12.15.: U-Test: Erkennungsleistung auf CVC-Ebene der Gruppen LBD und HC

| | W_sE | W_ges_Proz |
|-------------------------------------------|-------------------|-------------------|
| Mann-Whitney-U | 20,000 | 17,000 |
| Wilcoxon-W | 65,000 | 62,000 |
| Z | -2,459 | -2,752 |
| Asymptotische Signifikanz (2-seitig) | ,014 | ,006 |
| Exakte Signifikanz [2*(1-seitig Sig.)] | ,025 ^a | ,012 ^a |

a. Nicht für Bindungen korrigiert.

Abbildung 12.16.: U-Test: Erkennungsleistung auf W-Ebene der Gruppen LBD und RBD

12. Tabellen und Abbildungen

| | W sE |
|----------------------------------------|-------------------|
| Mann-Whitney-U | 57,000 |
| Wilcoxon-W | 210,000 |
| Z | -1,949 |
| Asymptotische Signifikanz (2-seitig) | ,051 |
| Exakte Signifikanz [2*(1-seitig Sig.)] | ,091 ^a |

a. Nicht für Bindungen korrigiert.

Abbildung 12.17.: U-Test: Erkennungsleistung auf W-Ebene der Gruppen LBD und HC

| | W ges. Proz | W vE |
|----------------------------------------|-------------------|-------------------|
| Mann-Whitney-U | 38,000 | 54,000 |
| Wilcoxon-W | 83,000 | 99,000 |
| Z | -2,284 | -1,775 |
| Asymptotische Signifikanz (2-seitig) | ,022 | ,076 |
| Exakte Signifikanz [2*(1-seitig Sig.)] | ,039 ^a | ,241 ^a |

a. Nicht für Bindungen korrigiert.

Abbildung 12.18.: U-Test: Erkennungsleistung auf W-Ebene der Gruppen RBD und HC

| | N | Korrelation | Signifikanz |
|--------------------------------|----|-------------|-------------|
| Paaren 1 CVC1_Proz & CVC2_Proz | 37 | ,014 | ,936 |
| Paaren 2 W1_Proz & W2_Proz | 37 | ,137 | ,419 |

Abbildung 12.19.: Korrelation zwischen Dauer und *percentage correct (PC)*

| | | Gepaarte Differenzen | | | | T | df | Sig. (2-seitig) | |
|----------|-----------------------|----------------------|--------------------|---------------------------------|--------------------------------------|----------|--------|-----------------|-------|
| | | Mittelwert | Standardabweichung | Standardfehler des Mittelwertes | 95% Konfidenzintervall der Differenz | | | | |
| | | | | | Untere | | | | Obere |
| Paaren 1 | CVC1_Proz - CVC2_Proz | 18,91892 | 65,98753 | 10,84828 | -3,08242 | 40,92026 | 1,744 | 36 | ,090 |
| Paaren 2 | W1_Proz - W2_Proz | -18,91892 | 56,94948 | 9,36244 | -37,90682 | ,06898 | -2,021 | 36 | ,051 |

Abbildung 12.20.: *percentage correct (PC)* der zeitkritischen Stimuli

12. Tabellen und Abbildungen

| | | Alter | Monate_po | HADS_D_Rohwert | AKT_T_PR | AKT_R_PR | AKT_F_PR | AKT_G_PR | WMS_R_Rohwert | PC_passed |
|----------------|--------------------------|-------|-----------|----------------|----------|----------|----------|----------|---------------|-----------|
| Alter | Korrelation nach Pearson | 1 | ,115 | -,014 | -,017 | -,538 | ^ | -,538 | ,420 | -,405 |
| | Signifikanz (2-seitig) | | ,736 | ,967 | ,961 | ,088 | ^ | ,088 | ,198 | ,216 |
| | N | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |
| Monate_po | Korrelation nach Pearson | ,115 | 1 | ,492 | ,066 | ,080 | ^ | ,087 | ,192 | ,094 |
| | Signifikanz (2-seitig) | ,736 | | ,125 | ,848 | ,815 | ^ | ,799 | ,571 | ,783 |
| | N | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |
| HADS_D_Rohwert | Korrelation nach Pearson | -,014 | ,492 | 1 | ,057 | -,065 | ^ | -,056 | -,332 | ,296 |
| | Signifikanz (2-seitig) | ,967 | ,125 | | ,868 | ,849 | ^ | ,871 | ,318 | ,378 |
| | N | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |
| AKT_T_PR | Korrelation nach Pearson | -,017 | ,066 | ,057 | 1 | ,037 | ^ | ,046 | ,429 | ,471 |
| | Signifikanz (2-seitig) | ,961 | ,848 | ,868 | | ,914 | ^ | ,892 | ,187 | ,144 |
| | N | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |
| AKT_R_PR | Korrelation nach Pearson | -,538 | ,080 | -,065 | ,037 | 1 | ^ | 1,000** | -,170 | ,426 |
| | Signifikanz (2-seitig) | ,088 | ,815 | ,849 | ,914 | | ^ | ,000 | ,616 | ,191 |
| | N | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |
| AKT_F_PR | Korrelation nach Pearson | ^ | ^ | ^ | ^ | ^ | 1 | ^ | ^ | ^ |
| | Signifikanz (2-seitig) | ^ | ^ | ^ | ^ | ^ | | ^ | ^ | ^ |
| | N | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |
| AKT_G_PR | Korrelation nach Pearson | -,538 | ,087 | -,056 | ,046 | 1,000** | ^ | 1 | -,160 | ,430 |
| | Signifikanz (2-seitig) | ,088 | ,799 | ,871 | ,892 | ,000 | ^ | | ,639 | ,187 |
| | N | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |
| WMS_R_Rohwert | Korrelation nach Pearson | ,420 | ,192 | -,332 | ,429 | -,170 | ^ | 1 | 1 | -,395 |
| | Signifikanz (2-seitig) | ,198 | ,571 | ,318 | ,187 | ,616 | ^ | | ,639 | ,230 |
| | N | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |
| PC_passed | Korrelation nach Pearson | -,405 | ,094 | ,296 | ,471 | ,426 | ^ | ,430 | 1 | 1 |
| | Signifikanz (2-seitig) | ,216 | ,783 | ,378 | ,144 | ,191 | ^ | ,187 | ,230 | |
| | N | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

^ Kann nicht berechnet werden, da mindestens eine der Variablen konstant ist.

Abbildung 12.21.: Korrelationen zwischen den Vortests und der Erkennungsleistung (LBD)

| | | Alter | Monate_po | HADS_D_Rohwert | AKT_T_PR | AKT_R_PR | AKT_F_PR | AKT_G_PR | WMS_R_Rohwert | PC_passed |
|----------------|--------------------------|--------|-----------|----------------|----------|----------|----------|----------|---------------|-----------|
| Alter | Korrelation nach Pearson | 1 | ,000 | -,299 | -,713* | ,037 | -,717* | -,531 | ,484 | -,038 |
| | Signifikanz (2-seitig) | | ,999 | ,435 | ,031 | ,925 | ,030 | ,142 | ,187 | ,922 |
| | N | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |
| Monate_po | Korrelation nach Pearson | ,000 | 1 | -,397 | ,348 | -,756* | -,130 | -,403 | ,339 | ,729* |
| | Signifikanz (2-seitig) | ,999 | | ,290 | ,358 | ,018 | ,738 | ,282 | ,373 | ,026 |
| | N | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |
| HADS_D_Rohwert | Korrelation nach Pearson | -,299 | -,397 | 1 | ,237 | ,615 | ,541 | ,584 | -,382 | -,220 |
| | Signifikanz (2-seitig) | ,435 | ,290 | | ,539 | ,078 | ,133 | ,099 | ,310 | ,569 |
| | N | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |
| AKT_T_PR | Korrelation nach Pearson | -,713* | ,348 | ,237 | 1 | -,342 | ,363 | ,127 | -,100 | ,566 |
| | Signifikanz (2-seitig) | ,031 | ,358 | ,539 | | ,368 | ,336 | ,744 | ,797 | ,112 |
| | N | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |
| AKT_R_PR | Korrelation nach Pearson | ,037 | -,756* | ,615 | -,342 | 1 | ,238 | ,524 | -,418 | -,661 |
| | Signifikanz (2-seitig) | ,925 | ,018 | ,078 | ,368 | | ,538 | ,148 | ,263 | ,053 |
| | N | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |
| AKT_F_PR | Korrelation nach Pearson | -,717* | -,130 | ,541 | ,363 | ,238 | 1 | ,923** | -,633 | -,071 |
| | Signifikanz (2-seitig) | ,030 | ,738 | ,133 | ,336 | ,538 | | ,000 | ,067 | ,866 |
| | N | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |
| AKT_G_PR | Korrelation nach Pearson | -,531 | -,403 | ,584 | ,127 | ,524 | ,923** | 1 | -,693* | -,232 |
| | Signifikanz (2-seitig) | ,142 | ,282 | ,099 | ,744 | ,148 | ,000 | | ,039 | ,549 |
| | N | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |
| WMS_R_Rohwert | Korrelation nach Pearson | ,484 | ,339 | -,382 | -,100 | -,418 | -,633 | -,693* | 1 | ,091 |
| | Signifikanz (2-seitig) | ,187 | ,373 | ,310 | ,797 | ,263 | ,067 | ,039 | | ,816 |
| | N | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |
| PC_passed | Korrelation nach Pearson | -,038 | ,729* | -,220 | ,566 | -,661 | -,071 | -,232 | ,091 | 1 |
| | Signifikanz (2-seitig) | ,922 | ,026 | ,569 | ,112 | ,053 | ,866 | ,549 | ,816 | |
| | N | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |

* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Abbildung 12.22.: Korrelationen zwischen den Vortests und der Erkennungsleistung (RBD)

12. Tabellen und Abbildungen

| | | Alter | HADS_D_Rohwert | AKT_T_PR | AKT_R_PR | AKT_F_PR | AKT_G_PR | WMS_R_Rohwert | PC_passed |
|----------------|--------------------------|-------|----------------|----------|----------|----------|----------|---------------|-----------|
| Alter | Korrelation nach Pearson | 1 | ,502* | ,423 | -,095 | -,081 | -,108 | ,012 | -,139 |
| | Signifikanz (2-seitig) | | ,040 | ,091 | ,718 | ,758 | ,680 | ,965 | ,594 |
| | N | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 |
| HADS_D_Rohwert | Korrelation nach Pearson | ,502* | 1 | -,027 | -,140 | -,454 | -,268 | ,289 | -,039 |
| | Signifikanz (2-seitig) | ,040 | | ,919 | ,592 | ,067 | ,299 | ,261 | ,882 |
| | N | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 |
| AKT_T_PR | Korrelation nach Pearson | ,423 | -,027 | 1 | ,044 | ,412 | ,167 | ,133 | -,141 |
| | Signifikanz (2-seitig) | ,091 | ,919 | | ,867 | ,101 | ,521 | ,611 | ,589 |
| | N | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 |
| AKT_R_PR | Korrelation nach Pearson | -,095 | -,140 | ,044 | 1 | ,228 | ,952** | ,475 | -,109 |
| | Signifikanz (2-seitig) | ,718 | ,592 | ,867 | | ,379 | ,000 | ,054 | ,676 |
| | N | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 |
| AKT_F_PR | Korrelation nach Pearson | -,081 | -,454 | ,412 | ,228 | 1 | ,515* | ,115 | ,009 |
| | Signifikanz (2-seitig) | ,758 | ,067 | ,101 | ,379 | | ,034 | ,661 | ,973 |
| | N | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 |
| AKT_G_PR | Korrelation nach Pearson | -,108 | -,268 | ,167 | ,952** | ,515* | 1 | ,454 | -,091 |
| | Signifikanz (2-seitig) | ,680 | ,299 | ,521 | ,000 | ,034 | | ,067 | ,730 |
| | N | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 |
| WMS_R_Rohwert | Korrelation nach Pearson | ,012 | ,289 | ,133 | ,475 | ,115 | ,454 | 1 | ,288 |
| | Signifikanz (2-seitig) | ,965 | ,261 | ,611 | ,054 | ,661 | ,067 | | ,263 |
| | N | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 |
| PC_passed | Korrelation nach Pearson | -,139 | -,039 | -,141 | -,109 | ,009 | -,091 | ,288 | 1 |
| | Signifikanz (2-seitig) | ,594 | ,882 | ,589 | ,676 | ,973 | ,730 | ,263 | |
| | N | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 |

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

**.. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Abbildung 12.23.: Korrelationen zwischen den Vortests und der Erkennungsleistung (HC)

12. Tabellen und Abbildungen

| | | Alter | Monate_po | HADS_D_Rohwert | AKT_T_PR | AKT_R_PR | AKT_F_PR | AKT_G_PR | WMS_R_Rohwert | PC_passed |
|----------------|--------------------------|--------|-----------|----------------|----------|----------|----------|----------|---------------|-----------|
| Alter | Korrelation nach Pearson | 1 | ,099 | -,167 | -,452* | -,115 | -,467* | -,477* | ,414 | -,087 |
| | Signifikanz (2-seitig) | | ,677 | ,483 | ,046 | ,631 | ,038 | ,033 | ,069 | ,714 |
| | N | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| Monate_po | Korrelation nach Pearson | ,099 | 1 | ,269 | ,170 | ,072 | ,085 | ,047 | ,116 | ,313 |
| | Signifikanz (2-seitig) | ,677 | | ,252 | ,473 | ,762 | ,723 | ,843 | ,626 | ,179 |
| | N | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| HADS_D_Rohwert | Korrelation nach Pearson | -,167 | ,269 | 1 | ,153 | ,230 | ,331 | ,283 | -,345 | ,019 |
| | Signifikanz (2-seitig) | ,483 | ,252 | | ,519 | ,330 | ,154 | ,226 | ,136 | ,935 |
| | N | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| AKT_T_PR | Korrelation nach Pearson | -,452* | ,170 | ,153 | 1 | -,046 | ,387 | ,166 | ,015 | ,561* |
| | Signifikanz (2-seitig) | ,046 | ,473 | ,519 | | ,846 | ,091 | ,484 | ,951 | ,010 |
| | N | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| AKT_R_PR | Korrelation nach Pearson | -,115 | ,072 | ,230 | -,046 | 1 | ,339 | ,726** | -,363 | ,157 |
| | Signifikanz (2-seitig) | ,631 | ,762 | ,330 | ,846 | | ,144 | ,000 | ,116 | ,507 |
| | N | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| AKT_F_PR | Korrelation nach Pearson | -,467* | ,085 | ,331 | ,387 | ,339 | 1 | ,769** | -,495* | ,183 |
| | Signifikanz (2-seitig) | ,038 | ,723 | ,154 | ,091 | ,144 | | ,000 | ,026 | ,441 |
| | N | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| AKT_G_PR | Korrelation nach Pearson | -,477* | ,047 | ,283 | ,166 | ,726** | ,769** | 1 | -,500* | ,170 |
| | Signifikanz (2-seitig) | ,033 | ,843 | ,226 | ,484 | ,000 | ,000 | | ,025 | ,473 |
| | N | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| WMS_R_Rohwert | Korrelation nach Pearson | ,414 | ,116 | -,345 | ,015 | -,363 | -,495* | -,500* | 1 | -,228 |
| | Signifikanz (2-seitig) | ,069 | ,626 | ,136 | ,951 | ,116 | ,026 | ,025 | | ,335 |
| | N | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| PC_passed | Korrelation nach Pearson | -,087 | ,313 | ,019 | ,561* | ,157 | ,183 | ,170 | -,228 | 1 |
| | Signifikanz (2-seitig) | ,714 | ,179 | ,935 | ,010 | ,507 | ,441 | ,473 | ,335 | |
| | N | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Abbildung 12.24.: Korrelationen zwischen den Vortests und der Erkennungsleistung (BD)

12. Tabellen und Abbildungen

| | | Geschlecht | Alter | Monate_po | HADS_D_Rohwert | AKT_T_PR | AKT_R_PR | AKT_F_PR | AKT_G_PR | WMS_R_Rohwert | PC_passed |
|----------------|--------------------------|------------|--------|-----------|----------------|----------|----------|----------|----------|---------------|-----------|
| Geschlecht | Korrelation nach Pearson | 1 | ,133 | -,341 | ,219 | -,192 | -,017 | ,216 | ,103 | -,073 | -,183 |
| | Signifikanz (2-seitig) | | ,432 | ,141 | ,193 | ,255 | ,922 | ,199 | ,546 | ,666 | ,278 |
| | N | 37 | 37 | 20 | 37 | 37 | 37 | 37 | 37 | 37 | 37 |
| Alter | Korrelation nach Pearson | ,133 | 1 | ,099 | ,040 | -,243 | -,120 | -,369* | -,384* | ,300 | -,104 |
| | Signifikanz (2-seitig) | ,432 | | ,677 | ,813 | ,148 | ,479 | ,025 | ,019 | ,539 | ,539 |
| | N | 37 | 37 | 20 | 37 | 37 | 37 | 37 | 37 | 37 | 37 |
| Monate_po | Korrelation nach Pearson | -,341 | ,099 | 1 | ,269 | ,170 | ,072 | ,085 | ,047 | ,116 | ,313 |
| | Signifikanz (2-seitig) | ,141 | ,677 | | ,252 | ,473 | ,762 | ,723 | ,843 | ,626 | ,179 |
| | N | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| HADS_D_Rohwert | Korrelation nach Pearson | ,219 | ,040 | ,269 | 1 | ,049 | ,047 | ,040 | ,065 | -,144 | -,015 |
| | Signifikanz (2-seitig) | ,193 | ,813 | ,252 | | ,772 | ,783 | ,814 | ,702 | ,394 | ,929 |
| | N | 37 | 37 | 20 | 37 | 37 | 37 | 37 | 37 | 37 | 37 |
| AKT_T_PR | Korrelation nach Pearson | -,192 | -,243 | ,170 | ,049 | 1 | ,018 | ,404* | ,185 | ,063 | ,301 |
| | Signifikanz (2-seitig) | ,255 | ,148 | ,473 | ,772 | | ,914 | ,013 | ,274 | ,710 | ,070 |
| | N | 37 | 37 | 20 | 37 | 37 | 37 | 37 | 37 | 37 | 37 |
| AKT_R_PR | Korrelation nach Pearson | -,017 | -,120 | ,072 | ,047 | ,018 | 1 | ,312 | ,805** | -,060 | ,055 |
| | Signifikanz (2-seitig) | ,922 | ,479 | ,762 | ,783 | ,914 | | ,060 | ,000 | ,725 | ,748 |
| | N | 37 | 37 | 20 | 37 | 37 | 37 | 37 | 37 | 37 | 37 |
| AKT_F_PR | Korrelation nach Pearson | ,216 | -,369* | ,085 | ,040 | ,404* | ,312 | 1 | ,692** | -,295 | ,119 |
| | Signifikanz (2-seitig) | ,199 | ,025 | ,723 | ,814 | ,013 | ,060 | | ,000 | ,000 | ,077 |
| | N | 37 | 37 | 20 | 37 | 37 | 37 | 37 | 37 | 37 | 37 |
| AKT_G_PR | Korrelation nach Pearson | ,103 | -,384* | ,047 | ,065 | ,185 | ,805** | ,692** | 1 | -,190 | ,075 |
| | Signifikanz (2-seitig) | ,546 | ,019 | ,843 | ,702 | ,274 | ,000 | ,000 | | ,260 | ,660 |
| | N | 37 | 37 | 20 | 37 | 37 | 37 | 37 | 37 | 37 | 37 |
| WMS_R_Rohwert | Korrelation nach Pearson | -,073 | ,300 | ,116 | -,144 | ,063 | -,060 | -,295 | -,190 | 1 | -,028 |
| | Signifikanz (2-seitig) | ,666 | ,071 | ,626 | ,394 | ,710 | ,725 | ,077 | ,260 | | ,871 |
| | N | 37 | 37 | 20 | 37 | 37 | 37 | 37 | 37 | 37 | 37 |
| PC_passed | Korrelation nach Pearson | -,183 | -,104 | ,313 | -,015 | ,301 | ,055 | ,119 | ,075 | -,028 | 1 |
| | Signifikanz (2-seitig) | ,278 | ,539 | ,179 | ,929 | ,070 | ,748 | ,483 | ,660 | ,871 | |
| | N | 37 | 37 | 20 | 37 | 37 | 37 | 37 | 37 | 37 | 37 |

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Abbildung 12.25.: Korrelationen zwischen den Vortests und der Erkennungsleistung (Gesamt)

12. Tabellen und Abbildungen

| | | | V1_Erk | V2_Erk | Bew_V1_schwierig | Bew_V1_sicher | Bew_V2_schwierig | Bew_V2_sicher |
|--------------|------------------|-------------------------|--------|--------|------------------|---------------|------------------|---------------|
| Spearman-Rho | V1_Erk | Korrelationskoeffizient | 1,000 | ,261 | ,000 | ,418 | -,407 | -,171 |
| | | Sig. (2-seitig) | . | ,438 | 1,000 | ,201 | ,214 | ,615 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | V2_Erk | Korrelationskoeffizient | ,261 | 1,000 | ,034 | -,187 | ,577 | ,581 |
| | | Sig. (2-seitig) | ,438 | . | ,922 | ,582 | ,063 | ,061 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_V1_schwierig | Korrelationskoeffizient | ,000 | ,034 | 1,000 | ,641* | ,217 | ,025 |
| | | Sig. (2-seitig) | 1,000 | ,922 | . | ,034 | ,522 | ,941 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_V1_sicher | Korrelationskoeffizient | ,418 | -,187 | ,641* | 1,000 | -,101 | -,384 |
| | | Sig. (2-seitig) | ,201 | ,582 | ,034 | . | ,769 | ,243 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_V2_schwierig | Korrelationskoeffizient | -,407 | ,577 | ,217 | -,101 | 1,000 | ,772** |
| | | Sig. (2-seitig) | ,214 | ,063 | ,522 | ,769 | . | ,005 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_V2_sicher | Korrelationskoeffizient | -,171 | ,581 | ,025 | -,384 | ,772** | 1,000 |
| | | Sig. (2-seitig) | ,615 | ,061 | ,941 | ,243 | ,005 | . |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |

*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

**.. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Abbildung 12.26.: Korrelationen der Ratingskalen (LBD, V)

| | | | CVC1_Erk | CVC2_Erk | Bew_CVC1_schwierig | Bew_CVC1_sicher | Bew_CVC2_schwierig | Bew_CVC2_sicher |
|--------------|--------------------|-------------------------|----------|----------|--------------------|-----------------|--------------------|-----------------|
| Spearman-Rho | CVC1_Erk | Korrelationskoeffizient | 1,000 | -,194 | ,108 | -,064 | -,263 | ,055 |
| | | Sig. (2-seitig) | . | ,568 | ,752 | ,852 | ,435 | ,873 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | CVC2_Erk | Korrelationskoeffizient | -,194 | 1,000 | -,418 | ,288 | ,238 | -,177 |
| | | Sig. (2-seitig) | ,568 | . | ,201 | ,390 | ,482 | ,603 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_CVC1_schwierig | Korrelationskoeffizient | ,108 | -,418 | 1,000 | ,322 | ,062 | -,210 |
| | | Sig. (2-seitig) | ,752 | ,201 | . | ,334 | ,857 | ,536 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_CVC1_sicher | Korrelationskoeffizient | -,064 | ,288 | ,322 | 1,000 | ,119 | ,060 |
| | | Sig. (2-seitig) | ,852 | ,390 | ,334 | . | ,728 | ,860 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_CVC2_schwierig | Korrelationskoeffizient | -,263 | ,238 | ,062 | ,119 | 1,000 | ,414 |
| | | Sig. (2-seitig) | ,435 | ,482 | ,857 | ,728 | . | ,206 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_CVC2_sicher | Korrelationskoeffizient | ,055 | -,177 | -,210 | ,060 | ,414 | 1,000 |
| | | Sig. (2-seitig) | ,873 | ,603 | ,536 | ,860 | ,206 | . |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |

Abbildung 12.27.: Korrelationen der Ratingskalen (LBD, CVC)

| | | | W1_Erk | W2_Erk | Bew_W1_schwierig | Bew_W1_sicher | Bew_W2_schwierig | Bew_W2_sicher |
|--------------|------------------|-------------------------|--------|--------|------------------|---------------|------------------|---------------|
| Spearman-Rho | W1_Erk | Korrelationskoeffizient | 1,000 | -,100 | ,289 | ,000 | ,162 | ,221 |
| | | Sig. (2-seitig) | . | ,770 | ,389 | 1,000 | ,635 | ,514 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | W2_Erk | Korrelationskoeffizient | -,100 | 1,000 | -,346 | ,000 | ,486 | ,553 |
| | | Sig. (2-seitig) | ,770 | . | ,297 | 1,000 | ,130 | ,078 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_W1_schwierig | Korrelationskoeffizient | ,289 | -,346 | 1,000 | ,247 | ,187 | -,191 |
| | | Sig. (2-seitig) | ,389 | ,297 | . | ,464 | ,582 | ,573 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_W1_sicher | Korrelationskoeffizient | ,000 | ,000 | ,247 | 1,000 | ,189 | ,430 |
| | | Sig. (2-seitig) | 1,000 | 1,000 | ,464 | . | ,578 | ,186 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_W2_schwierig | Korrelationskoeffizient | ,162 | ,486 | ,187 | ,189 | 1,000 | ,716* |
| | | Sig. (2-seitig) | ,635 | ,130 | ,582 | ,578 | . | ,013 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_W2_sicher | Korrelationskoeffizient | ,221 | ,553 | -,191 | ,430 | ,716* | 1,000 |
| | | Sig. (2-seitig) | ,514 | ,078 | ,573 | ,186 | ,013 | . |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |

*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Abbildung 12.28.: Korrelationen der Ratingskalen (LBD, W)

12. Tabellen und Abbildungen

| | | | S1_Erk | S2_Erk | Bew_S1_ schwierig | Bew_S1_ sicher | Bew_S2_ schwierig | Bew_S2_ sicher |
|--------------|------------------|-------------------------|--------|--------|----------------------|-------------------|----------------------|-------------------|
| Spearman-Rho | S1_Erk | Korrelationskoeffizient | .. | .. | .. | .. | .. | .. |
| | | Sig. (2-seitig) | .. | .. | .. | .. | .. | .. |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | S2_Erk | Korrelationskoeffizient | .. | .. | .. | .. | .. | .. |
| | | Sig. (2-seitig) | .. | .. | .. | .. | .. | .. |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_S1_schwierig | Korrelationskoeffizient | .. | .. | 1,000 | ,671* | 1,000** | ,671* |
| | | Sig. (2-seitig) | .. | .. | .. | ,024 | .. | ,024 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_S1_sicher | Korrelationskoeffizient | .. | .. | ,671* | 1,000 | ,671* | 1,000** |
| | | Sig. (2-seitig) | .. | .. | ,024 | .. | ,024 | .. |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_S2_schwierig | Korrelationskoeffizient | .. | .. | 1,000** | ,671* | 1,000 | ,671* |
| | | Sig. (2-seitig) | .. | .. | .. | ,024 | .. | ,024 |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |
| | Bew_S2_sicher | Korrelationskoeffizient | .. | .. | ,671* | 1,000** | ,671* | 1,000 |
| | | Sig. (2-seitig) | .. | .. | ,024 | .. | ,024 | .. |
| | | N | 11 | 11 | 11 | 11 | 11 | 11 |

*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

**.. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Abbildung 12.29.: Korrelationen der Ratingskalen (LBD, S)

12. Tabellen und Abbildungen

| | | | V1_Erk | V2_Erk | Bew_V1_schwierig | Bew_V1_sicher | Bew_V2_schwierig | Bew_V2_sicher |
|--------------|------------------|-------------------------|--------|--------|------------------|---------------|------------------|---------------|
| Spearman-Rho | V1_Erk | Korrelationskoeffizient | 1,000 | ,316 | ,880** | ,755* | ,867** | ,636 |
| | | Sig. (2-seitig) | . | ,407 | ,002 | ,019 | ,002 | ,066 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | V2_Erk | Korrelationskoeffizient | ,316 | 1,000 | ,342 | ,234 | ,385 | ,527 |
| | | Sig. (2-seitig) | ,407 | . | ,368 | ,544 | ,306 | ,145 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_V1_schwierig | Korrelationskoeffizient | ,880** | ,342 | 1,000 | ,667* | ,986** | ,566 |
| | | Sig. (2-seitig) | ,002 | ,368 | . | ,050 | ,000 | ,113 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_V1_sicher | Korrelationskoeffizient | ,755* | ,234 | ,667* | 1,000 | ,707* | ,884** |
| | | Sig. (2-seitig) | ,019 | ,544 | ,050 | . | ,033 | ,002 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_V2_schwierig | Korrelationskoeffizient | ,867** | ,385 | ,986** | ,707* | 1,000 | ,618 |
| | | Sig. (2-seitig) | ,002 | ,306 | ,000 | ,033 | . | ,076 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_V2_sicher | Korrelationskoeffizient | ,636 | ,527 | ,566 | ,884** | ,618 | 1,000 |
| | | Sig. (2-seitig) | ,066 | ,145 | ,113 | ,002 | ,076 | . |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |

** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

* Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Abbildung 12.30.: Korrelationen der Ratingskalen (RBD, V)

| | | | CVC1_Erk | CVC2_Erk | Bew_CVC1_schwierig | Bew_CVC1_sicher | Bew_CVC2_schwierig | Bew_CVC2_sicher |
|--------------|--------------------|-------------------------|----------|----------|--------------------|-----------------|--------------------|-----------------|
| Spearman-Rho | CVC1_Erk | Korrelationskoeffizient | 1,000 | -,060 | ,265 | ,441 | ,267 | ,330 |
| | | Sig. (2-seitig) | . | ,879 | ,490 | ,235 | ,488 | ,385 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | CVC2_Erk | Korrelationskoeffizient | -,060 | 1,000 | ,400 | -,046 | ,357 | -,138 |
| | | Sig. (2-seitig) | ,879 | . | ,286 | ,906 | ,346 | ,723 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_CVC1_schwierig | Korrelationskoeffizient | ,265 | ,400 | 1,000 | ,428 | ,907** | ,264 |
| | | Sig. (2-seitig) | ,490 | ,286 | . | ,251 | ,001 | ,493 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_CVC1_sicher | Korrelationskoeffizient | ,441 | -,046 | ,428 | 1,000 | ,247 | ,443 |
| | | Sig. (2-seitig) | ,235 | ,906 | ,251 | . | ,522 | ,232 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_CVC2_schwierig | Korrelationskoeffizient | ,267 | ,357 | ,907** | ,247 | 1,000 | ,484 |
| | | Sig. (2-seitig) | ,488 | ,346 | ,001 | ,522 | . | ,186 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_CVC2_sicher | Korrelationskoeffizient | ,330 | -,138 | ,264 | ,443 | ,484 | 1,000 |
| | | Sig. (2-seitig) | ,385 | ,723 | ,493 | ,232 | ,186 | . |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |

** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Abbildung 12.31.: Korrelationen der Ratingskalen (RBD, CVC)

12. Tabellen und Abbildungen

| | | | W1_Erk | W2_Erk | Bew_W1_schwierig | Bew_W1_sicher | Bew_W2_schwierig | Bew_W2_sicher |
|--------------|------------------|-------------------------|--------|--------|------------------|---------------|------------------|---------------|
| Spearman-Rho | W1_Erk | Korrelationskoeffizient | 1,000 | ,158 | ,283 | ,732* | ,096 | ,439 |
| | | Sig. (2-seitig) | . | ,685 | ,460 | ,025 | ,805 | ,237 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | W2_Erk | Korrelationskoeffizient | ,158 | 1,000 | ,269 | ,463 | ,502 | ,463 |
| | | Sig. (2-seitig) | ,685 | . | ,484 | ,210 | ,168 | ,210 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_W1_schwierig | Korrelationskoeffizient | ,283 | ,269 | 1,000 | ,526 | ,900** | ,526 |
| | | Sig. (2-seitig) | ,460 | ,484 | . | ,146 | ,001 | ,146 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_W1_sicher | Korrelationskoeffizient | ,732* | ,463 | ,526 | 1,000 | ,578 | ,829** |
| | | Sig. (2-seitig) | ,025 | ,210 | ,146 | . | ,103 | ,006 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_W2_schwierig | Korrelationskoeffizient | ,096 | ,502 | ,900** | ,578 | 1,000 | ,690* |
| | | Sig. (2-seitig) | ,805 | ,168 | ,001 | ,103 | . | ,040 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_W2_sicher | Korrelationskoeffizient | ,439 | ,463 | ,526 | ,829** | ,690* | 1,000 |
| | | Sig. (2-seitig) | ,237 | ,210 | ,146 | ,006 | ,040 | . |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |

*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

**. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Abbildung 12.32.: Korrelationen der Ratingskalen (RBD, W)

| | | | S1_Erk | S2_Erk | Bew_S1_schwierig | Bew_S1_sicher | Bew_S2_schwierig | Bew_S2_sicher |
|--------------|------------------|-------------------------|--------|--------|------------------|---------------|------------------|---------------|
| Spearman-Rho | S1_Erk | Korrelationskoeffizient | . | . | . | . | . | . |
| | | Sig. (2-seitig) | . | . | . | . | . | . |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | S2_Erk | Korrelationskoeffizient | . | . | . | . | . | . |
| | | Sig. (2-seitig) | . | . | . | . | . | . |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_S1_schwierig | Korrelationskoeffizient | . | . | 1,000 | ,704* | ,895** | ,617 |
| | | Sig. (2-seitig) | . | . | . | ,034 | ,001 | ,077 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_S1_sicher | Korrelationskoeffizient | . | . | ,704* | 1,000 | ,609 | ,696* |
| | | Sig. (2-seitig) | . | . | ,034 | . | ,082 | ,037 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_S2_schwierig | Korrelationskoeffizient | . | . | ,895** | ,609 | 1,000 | ,686* |
| | | Sig. (2-seitig) | . | . | ,001 | ,082 | . | ,041 |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |
| | Bew_S2_sicher | Korrelationskoeffizient | . | . | ,617 | ,696* | ,686* | 1,000 |
| | | Sig. (2-seitig) | . | . | ,077 | ,037 | ,041 | . |
| | | N | 9 | 9 | 9 | 9 | 9 | 9 |

*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

**. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Abbildung 12.33.: Korrelationen der Ratingskalen (RBD, S)

12. Tabellen und Abbildungen

| | | | V1_Erk | V2_Erk | Bew_V1_schwierig | Bew_V1_sicher | Bew_V2_schwierig | Bew_V2_sicher |
|--------------|------------------|-------------------------|--------|--------|------------------|---------------|------------------|---------------|
| Spearman-Rho | V1_Erk | Korrelationskoeffizient | 1,000 | ,588* | -,179 | ,321 | -,472 | -,270 |
| | | Sig. (2-seitig) | . | ,013 | ,492 | ,209 | ,056 | ,294 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | V2_Erk | Korrelationskoeffizient | ,588* | 1,000 | ,266 | ,372 | -,125 | ,128 |
| | | Sig. (2-seitig) | ,013 | . | ,302 | ,142 | ,631 | ,626 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_V1_schwierig | Korrelationskoeffizient | -,179 | ,266 | 1,000 | ,615** | ,754** | ,696** |
| | | Sig. (2-seitig) | ,492 | ,302 | . | ,009 | ,000 | ,002 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_V1_sicher | Korrelationskoeffizient | ,321 | ,372 | ,615** | 1,000 | ,248 | ,533* |
| | | Sig. (2-seitig) | ,209 | ,142 | ,009 | . | ,338 | ,027 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_V2_schwierig | Korrelationskoeffizient | -,472 | -,125 | ,754** | ,248 | 1,000 | ,765** |
| | | Sig. (2-seitig) | ,056 | ,631 | ,000 | ,338 | . | ,000 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_V2_sicher | Korrelationskoeffizient | -,270 | ,128 | ,696** | ,533* | ,765** | 1,000 |
| | | Sig. (2-seitig) | ,294 | ,626 | ,002 | ,027 | ,000 | . |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |

*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

**.. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Abbildung 12.34.: Korrelationen der Ratingskalen (HC, V)

| | | | CVC1_Erk | CVC2_Erk | Bew_CVC1_schwierig | Bew_CVC1_sicher | Bew_CVC2_schwierig | Bew_CVC2_sicher |
|--------------|--------------------|-------------------------|----------|----------|--------------------|-----------------|--------------------|-----------------|
| Spearman-Rho | CVC1_Erk | Korrelationskoeffizient | 1,000 | ,044 | ,393 | ,611** | -,206 | -,054 |
| | | Sig. (2-seitig) | . | ,868 | ,119 | ,009 | ,428 | ,837 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | CVC2_Erk | Korrelationskoeffizient | ,044 | 1,000 | ,246 | ,411 | ,394 | ,258 |
| | | Sig. (2-seitig) | ,868 | . | ,340 | ,101 | ,117 | ,317 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_CVC1_schwierig | Korrelationskoeffizient | ,393 | ,246 | 1,000 | ,757** | ,214 | ,473 |
| | | Sig. (2-seitig) | ,119 | ,340 | . | ,000 | ,410 | ,055 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_CVC1_sicher | Korrelationskoeffizient | ,611** | ,411 | ,757** | 1,000 | ,132 | ,353 |
| | | Sig. (2-seitig) | ,009 | ,101 | ,000 | . | ,613 | ,164 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_CVC2_schwierig | Korrelationskoeffizient | -,206 | ,394 | ,214 | ,132 | 1,000 | ,562* |
| | | Sig. (2-seitig) | ,428 | ,117 | ,410 | ,613 | . | ,019 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_CVC2_sicher | Korrelationskoeffizient | -,054 | ,258 | ,473 | ,353 | ,562* | 1,000 |
| | | Sig. (2-seitig) | ,837 | ,317 | ,055 | ,164 | ,019 | . |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |

**.. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Abbildung 12.35.: Korrelationen der Ratingskalen (HC, CVC)

12. Tabellen und Abbildungen

| | | | W1_Erk | W2_Erk | Bew_W1_ schwierig | Bew_W1_ sicher | Bew_W2_ schwierig | Bew_W2_ sicher |
|--------------|----------------------|-------------------------|--------|--------|----------------------|-------------------|----------------------|-------------------|
| Spearman-Rho | W1_Erk | Korrelationskoeffizient | 1,000 | -,185 | ,107 | ,143 | -,074 | ,027 |
| | | Sig. (2-seitig) | . | ,478 | ,684 | ,583 | ,778 | ,917 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | W2_Erk | Korrelationskoeffizient | -,185 | 1,000 | ,298 | ,344 | ,450 | ,414 |
| | | Sig. (2-seitig) | ,478 | . | ,246 | ,176 | ,070 | ,098 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_W1_ schwierig | Korrelationskoeffizient | ,107 | ,298 | 1,000 | ,811** | ,525* | ,337 |
| | | Sig. (2-seitig) | ,684 | ,246 | . | ,000 | ,030 | ,186 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_W1_ sicher | Korrelationskoeffizient | ,143 | ,344 | ,811** | 1,000 | ,589* | ,484* |
| | | Sig. (2-seitig) | ,583 | ,176 | ,000 | . | ,013 | ,049 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_W2_ schwierig | Korrelationskoeffizient | -,074 | ,450 | ,525* | ,589* | 1,000 | ,591* |
| | | Sig. (2-seitig) | ,778 | ,070 | ,030 | ,013 | . | ,012 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_W2_ sicher | Korrelationskoeffizient | ,027 | ,414 | ,337 | ,484* | ,591* | 1,000 |
| | | Sig. (2-seitig) | ,917 | ,098 | ,186 | ,049 | ,012 | . |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |

** . Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

* . Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Abbildung 12.36.: Korrelationen der Ratingskalen (HC, W)

| | | | S1_Erk | S2_Erk | Bew_S1_ schwierig | Bew_S1_ sicher | Bew_S2_ schwierig | Bew_S2_ sicher |
|--------------|----------------------|-------------------------|--------|--------|----------------------|-------------------|----------------------|-------------------|
| Spearman-Rho | S1_Erk | Korrelationskoeffizient | . | . | . | . | . | . |
| | | Sig. (2-seitig) | . | . | . | . | . | . |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | S2_Erk | Korrelationskoeffizient | . | . | . | . | . | . |
| | | Sig. (2-seitig) | . | . | . | . | . | . |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_S1_ schwierig | Korrelationskoeffizient | . | . | 1,000 | ,614** | ,757** | ,598* |
| | | Sig. (2-seitig) | . | . | . | ,009 | ,000 | ,011 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_S1_ sicher | Korrelationskoeffizient | . | . | ,614** | 1,000 | ,627** | ,993** |
| | | Sig. (2-seitig) | . | . | ,009 | . | ,007 | ,000 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_S2_ schwierig | Korrelationskoeffizient | . | . | ,757** | ,627** | 1,000 | ,639** |
| | | Sig. (2-seitig) | . | . | ,000 | ,007 | . | ,006 |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |
| | Bew_S2_ sicher | Korrelationskoeffizient | . | . | ,598* | ,993** | ,639** | 1,000 |
| | | Sig. (2-seitig) | . | . | ,011 | ,000 | ,006 | . |
| | | N | 17 | 17 | 17 | 17 | 17 | 17 |

** . Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

* . Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Abbildung 12.37.: Korrelationen der Ratingskalen (HC, S)

12. Tabellen und Abbildungen

| | | | V1_Erk | V2_Erk | Bew_V1_schwierig | Bew_V1_sicher | Bew_V2_schwierig | Bew_V2_sicher |
|--------------|------------------|-------------------------|--------|--------|------------------|---------------|------------------|---------------|
| Spearman-Rho | V1_Erk | Korrelationskoeffizient | 1,000 | ,443** | ,210 | ,509** | -,014 | ,073 |
| | | Sig. (2-seitig) | . | ,006 | ,212 | ,001 | ,934 | ,667 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | V2_Erk | Korrelationskoeffizient | ,443** | 1,000 | ,298 | ,266 | ,235 | ,383* |
| | | Sig. (2-seitig) | ,006 | . | ,073 | ,112 | ,161 | ,019 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_V1_schwierig | Korrelationskoeffizient | ,210 | ,298 | 1,000 | ,650** | ,704** | ,534** |
| | | Sig. (2-seitig) | ,212 | ,073 | . | ,000 | ,000 | ,001 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_V1_sicher | Korrelationskoeffizient | ,509** | ,266 | ,650** | 1,000 | ,340* | ,492** |
| | | Sig. (2-seitig) | ,001 | ,112 | ,000 | . | ,040 | ,002 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_V2_schwierig | Korrelationskoeffizient | -,014 | ,235 | ,704** | ,340* | 1,000 | ,732** |
| | | Sig. (2-seitig) | ,934 | ,161 | ,000 | ,040 | . | ,000 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_V2_sicher | Korrelationskoeffizient | ,073 | ,383* | ,534** | ,492** | ,732** | 1,000 |
| | | Sig. (2-seitig) | ,667 | ,019 | ,001 | ,002 | ,000 | . |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |

** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

* Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Abbildung 12.38.: Korrelationen der Ratingskalen (Gesamt, V)

| | | | CVC1_Erk | CVC2_Erk | Bew_CVC1_schwierig | Bew_CVC1_sicher | Bew_CVC2_schwierig | Bew_CVC2_sicher |
|--------------|--------------------|-------------------------|----------|----------|--------------------|-----------------|--------------------|-----------------|
| Spearman-Rho | CVC1_Erk | Korrelationskoeffizient | 1,000 | ,014 | ,278 | ,484** | -,059 | ,086 |
| | | Sig. (2-seitig) | . | ,936 | ,095 | ,002 | ,728 | ,614 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | CVC2_Erk | Korrelationskoeffizient | ,014 | 1,000 | ,141 | ,310 | ,334* | ,030 |
| | | Sig. (2-seitig) | ,936 | . | ,406 | ,062 | ,044 | ,859 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_CVC1_schwierig | Korrelationskoeffizient | ,278 | ,141 | 1,000 | ,593** | ,383* | ,193 |
| | | Sig. (2-seitig) | ,095 | ,406 | . | ,000 | ,019 | ,253 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_CVC1_sicher | Korrelationskoeffizient | ,484** | ,310 | ,593** | 1,000 | ,158 | ,296 |
| | | Sig. (2-seitig) | ,002 | ,062 | ,000 | . | ,351 | ,075 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_CVC2_schwierig | Korrelationskoeffizient | -,059 | ,334* | ,383* | ,158 | 1,000 | ,485** |
| | | Sig. (2-seitig) | ,728 | ,044 | ,019 | ,351 | . | ,002 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_CVC2_sicher | Korrelationskoeffizient | ,086 | ,030 | ,193 | ,296 | ,485** | 1,000 |
| | | Sig. (2-seitig) | ,614 | ,859 | ,253 | ,075 | ,002 | . |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |

** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

* Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Abbildung 12.39.: Korrelationen der Ratingskalen (Gesamt, CVC)

12. Tabellen und Abbildungen

| | | | W1_Erk | W2_Erk | Bew_W1_schwierig | Bew_W1_sicher | Bew_W2_schwierig | Bew_W2_sicher |
|--------------|------------------|-------------------------|--------|--------|------------------|---------------|------------------|---------------|
| Spearman-Rho | W1_Erk | Korrelationskoeffizient | 1,000 | ,137 | ,175 | ,253 | ,184 | ,238 |
| | | Sig. (2-seitig) | . | ,419 | ,301 | ,131 | ,276 | ,157 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | W2_Erk | Korrelationskoeffizient | ,137 | 1,000 | ,088 | ,155 | ,476** | ,486** |
| | | Sig. (2-seitig) | ,419 | . | ,606 | ,361 | ,003 | ,002 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_W1_schwierig | Korrelationskoeffizient | ,175 | ,088 | 1,000 | ,627** | ,576** | ,281 |
| | | Sig. (2-seitig) | ,301 | ,606 | . | ,000 | ,000 | ,092 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_W1_sicher | Korrelationskoeffizient | ,253 | ,155 | ,627** | 1,000 | ,462** | ,478** |
| | | Sig. (2-seitig) | ,131 | ,361 | ,000 | . | ,004 | ,003 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_W2_schwierig | Korrelationskoeffizient | ,184 | ,476** | ,576** | ,462** | 1,000 | ,712** |
| | | Sig. (2-seitig) | ,276 | ,003 | ,000 | ,004 | . | ,000 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_W2_sicher | Korrelationskoeffizient | ,238 | ,486** | ,281 | ,478** | ,712** | 1,000 |
| | | Sig. (2-seitig) | ,157 | ,002 | ,092 | ,003 | ,000 | . |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |

** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Abbildung 12.40.: Korrelationen der Ratingskalen (Gesamt, W)

| | | | S1_Erk | S2_Erk | Bew_S1_schwierig | Bew_S1_sicher | Bew_S2_schwierig | Bew_S2_sicher |
|--------------|------------------|-------------------------|--------|--------|------------------|---------------|------------------|---------------|
| Spearman-Rho | S1_Erk | Korrelationskoeffizient | . | . | . | . | . | . |
| | | Sig. (2-seitig) | . | . | . | . | . | . |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | S2_Erk | Korrelationskoeffizient | . | . | . | . | . | . |
| | | Sig. (2-seitig) | . | . | . | . | . | . |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_S1_schwierig | Korrelationskoeffizient | . | . | 1,000 | ,641** | ,834** | ,562** |
| | | Sig. (2-seitig) | . | . | . | ,000 | ,000 | ,000 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_S1_sicher | Korrelationskoeffizient | . | . | ,641** | 1,000 | ,661** | ,893** |
| | | Sig. (2-seitig) | . | . | ,000 | . | ,000 | ,000 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_S2_schwierig | Korrelationskoeffizient | . | . | ,834** | ,661** | 1,000 | ,640** |
| | | Sig. (2-seitig) | . | . | ,000 | ,000 | . | ,000 |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |
| | Bew_S2_sicher | Korrelationskoeffizient | . | . | ,562** | ,893** | ,640** | 1,000 |
| | | Sig. (2-seitig) | . | . | ,000 | ,000 | ,000 | . |
| | | N | 37 | 37 | 37 | 37 | 37 | 37 |

** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Abbildung 12.41.: Korrelationen der Ratingskalen (Gesamt, S)

13. Verwendete Test- und Untersuchungsbögen

13. Verwendete Test- und Untersuchungsbögen



Task Oriented Communication

Universität Bielefeld
Fakultät für Linguistik und Literaturwissenschaft
Graduiertenkolleg Aufgabenorientierte Kommunikation

Sehr geehrte Interessentin, sehr geehrter Interessent,

herzlichen Dank für Ihr Interesse an unserem Forschungsvorhaben!

Um die Funktions- und Arbeitsweise des menschlichen Gehirns noch besser verstehen zu können, führen wir zur Zeit innerhalb eines interdisziplinären Forschungsprojektes Untersuchungen mit Schlaganfall-Patienten an ausgewählten (Reha-)Kliniken in Ostwestfalen durch. Dabei sollen spezifische Leistungen des Gehirns untersucht und ausgewertet werden.

Im Rahmen eines Wahrnehmungs-Experimentes werden den Patienten unterschiedliche Stimmen vorgespielt, unter denen sie die eines nahen Angehörigen erkennen sollen.

Um nun eine ausreichende Menge an unbekannten („Ablenker-“)Stimmen zur Verfügung zu haben, sind wir auf Ihre Mithilfe angewiesen.

Ihre Aufgabe

Ihre Aufgabe besteht lediglich darin, vsl. acht Texttafeln vorzulesen. Dieser Vorgang dauert etwa 10 bis 20 Minuten und wird elektronisch gespeichert. Da hierfür keine besonderen Geräte erforderlich sind, kann die Aufnahme an jedem (ruhigen) beliebigen Ort (z.B. bei Ihnen zu Hause) durchgeführt werden.

Betreut wird die Studie von Frau Dr. Martina Hielscher-Fastabend und Herrn Prof. Dr. Gert Rickheit (beide Universität Bielefeld, Fakultät für Linguistik und Literaturwissenschaft) sowie von Herrn apl. Prof. Dr. med. Christoph Lang (Friedrich-Alexander-Universität Erlangen-Nürnberg, Neurologische Klinik).

Wir würden uns freuen, wenn Sie unser Forschungsprojekt durch Ihre Teilnahme unterstützen und die beiliegende Einverständniserklärung unterzeichnen würden.

Für etwaige Fragen stehen wir Ihnen selbstverständlich jederzeit gern zur Verfügung.

Mit freundlichen Grüßen

Oliver Kneidl (Tel. 0521 / 106-5324)

Datenschutz: Im Rahmen einer sinnvollen Auswertung benötigen wir von Ihnen lediglich die Angabe Ihres Alters. Andere Informationen oder persönliche Daten sind irrelevant und werden nicht erfragt!



Die Studie wird aus Mitteln der Deutschen Forschungsgemeinschaft (GK 256) finanziert.

Abbildung 13.1.: Informationsblatt für die Distraktoren bzw. Angehörigen

13. Verwendete Test- und Untersuchungsbögen



Task Oriented Communication

Universität Bielefeld
Fakultät für Linguistik und Literaturwissenschaft
Graduiertenkolleg Aufgabenorientierte Kommunikation

Sehr geehrte Interessentin, sehr geehrter Interessent,

herzlichen Dank für Ihr Interesse an unserem Forschungsvorhaben!

Um die Funktions- und Arbeitsweise des menschlichen Gehirns noch besser verstehen zu können, führen wir zur Zeit innerhalb eines interdisziplinären Forschungsprojektes Untersuchungen mit Schlaganfall-Patienten an ausgewählten (Reha-)Kliniken und logopädischen Praxen in Ostwestfalen durch.

Im Rahmen eines „Wahrnehmungs-Experimentes“ werden den Teilnehmer/inne/n unterschiedliche Stimmen vorgespielt, unter denen sie die eines nahen Angehörigen erkennen sollen.

Die Aufgabe des Patienten

Ihre Aufgabe besteht darin, aus fünf Stimmproben die Stimme Ihres Angehörigen herauszufinden. Insgesamt sind acht solcher Erkennungsaufgaben vorgesehen (Gesamtdauer: ca. 20-30 Minuten). Da u.U. noch einige Vortests nötig sind, erläutern wir Ihnen den genauen Ablauf gern in einem persönlichen Gespräch.

Die Aufgaben Ihres Angehörigen

Ihr/e Angehörige/r hat lediglich die Aufgabe, acht Sätze vorzulesen, die von uns aufgezeichnet werden (Gesamtdauer: max. 5-10 Minuten). Weitere Vortests sind hier nicht erforderlich.

Wir würden uns sehr freuen, wenn Sie unser Forschungsprojekt durch Ihre Teilnahme unterstützen würden. Je nach Möglichkeit kann das Experiment in der Praxis Ihres Logopäden, in Ihrer (Reha-)Klinik, bei Ihnen zu Hause oder bei uns an der Universität durchgeführt werden.

Für etwaige Fragen stehen wir Ihnen selbstverständlich jederzeit gern zur Verfügung.

Mit freundlichen Grüßen

Oliver Kneidl
(Tel. 0521 / 106-5324, Büro: D6-127, E-Mail: oliver.kneidl@uni-bielefeld.de)

Datenschutz: Alle im Rahmen der Untersuchung gemachten Angaben werden nur anonymisiert verwandt und unterliegen den datenschutzrechtlichen Bestimmungen. Alle beteiligten Personen unterliegen der Schweigepflicht.

Institutsangaben: Betreut wird die Studie von Frau Dr. Martina Hielscher-Fastabend und Herrn Prof. Dr. Gert Rickheit (beide Universität Bielefeld, Fakultät für Linguistik und Literaturwissenschaft) sowie von Herrn apl. Prof. Dr. med. Christoph Lang (Friedrich-Alexander-Universität Erlangen-Nürnberg, Neurologische Klinik).



Die Studie wird aus Mitteln der Deutschen Forschungsgemeinschaft (GK 256) finanziert.

Abbildung 13.2.: Informationsblatt für die Untersuchungsteilnehmer

13. Verwendete Test- und Untersuchungsbögen

Erfassungsbogen

| | | | | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|--|
| Name, Vorname | | Probandennr. | | |
| Geburtsdatum | | Geburtsort | | |
| Alter (jj;tt) | | Geschlecht <input type="checkbox"/> w (0) <input type="checkbox"/> m (1) | | |
| Läsion (nachgewiesen durch _____) <input type="checkbox"/> unilateral reH <input type="checkbox"/> unilateral liH <input type="checkbox"/> bilateral | | Erstinfarkt <input type="checkbox"/> Ja <input type="checkbox"/> Nein | Datum des Insultereignis | |
| Lokalisation <input type="checkbox"/> A. cerebri media <input type="checkbox"/> vordere Mediaastgruppe <input type="checkbox"/> mittlere Mediaastgruppe <input type="checkbox"/> hintere Mediaastgruppe <input type="checkbox"/> A. cerebri anterior <input type="checkbox"/> A. cerebri posterior | | Stationärer Behandlungszeitraum / AHB | | |
| | | Ambulante Versorgung (Logo) | | |
| | | Ätiologie <input type="checkbox"/> Infarkt <input type="checkbox"/> Tumor <input type="checkbox"/> SHT <input type="checkbox"/> _____ | | |
| Familienstand <input type="checkbox"/> ledig (1) <input type="checkbox"/> verheiratet (2) <input type="checkbox"/> geschieden (3) <input type="checkbox"/> getrennt lebend (4) <input type="checkbox"/> verwitwet (5) | | Schulabschluss <input type="checkbox"/> keinen / SoS (1) <input type="checkbox"/> VS / HS (2) <input type="checkbox"/> Mittl. Reife (3) <input type="checkbox"/> Abitur (4) <input type="checkbox"/> Hochschule (5) | | |
| Schulbildung (Dauer) | | Erlerner / ausgeübter Beruf | | |
| Händigkeit <input type="checkbox"/> links <input type="checkbox"/> rechts <input type="checkbox"/> ambidexter | | Muttersprache / Fremdsprache(n) | | |
| HNO <input type="checkbox"/> oB <input type="checkbox"/> Hörgerät/e <input type="checkbox"/> Hypakusis | | Zielstimme vorhanden <input type="checkbox"/> | | |

| AAT-Werte | Punktwerte | Prozentrang | Allgemein | Syndrom |
|---------------------------------------------------------------------|------------|-------------|-----------|---------|
| SPON | | | | |
| TT | | | | |
| NACH | | | | |
| SCHRIFT | | | | |
| BN | | | | |
| SV | | | | |
| AAT-Ergebnis (Diagnose, Spezifizierung) | | | | |
| MMST | | AKT | | |
| HADS-D | | WMS-R | | |
| V.a. FKHS <input type="checkbox"/> nein <input type="checkbox"/> ja | | Neglect | | |

Daten erfasst

Abbildung 13.3.: Anamnesebogen

13. Verwendete Test- und Untersuchungsbögen

| VERBALE PAARERKENNUNG I Wenn der Proband alle Aufgaben im 3. Durchgang löst, ist der Untertest abgeschlossen. Ansonsten führen Sie auch die folgenden Durchgänge durch, bis der Proband alle Aufgaben löst. | | | | | | | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|-----------------|------------------|---------------------|--------------|----------------|---------------|
| 1. Durchgang | Abruf | Einfach | Schwer | 4. Durchgang | Abruf | Einfach | Schwer |
| Metall - Eisen | Frucht | _____ | _____ | Gedränge - Dunkel | Schule | _____ | _____ |
| Baby - Geschrei | Gehorchen | _____ | _____ | Salat - Stift | Metall | _____ | _____ |
| Gedränge - Dunkel | Rose | _____ | _____ | Frucht - Apfel | Gehorchen | _____ | _____ |
| Schule - Bäckerei | Baby | _____ | _____ | Gehorchen - Meter | Gedränge | _____ | _____ |
| Rose - Blume | Salat | _____ | _____ | Baby - Geschrei | Frucht | _____ | _____ |
| Gehorchen - Meter | Metall | _____ | _____ | Rose - Blume | Baby | _____ | _____ |
| Frucht - Apfel | Schule | _____ | _____ | Metall - Eisen | Salat | _____ | _____ |
| Salat - Stift | Gedränge | _____ | _____ | Schule - Bäckerei | Rose | _____ | _____ |
| Gesamt | | _____ | _____ | Gesamt | | _____ | _____ |
| 2. Durchgang | Abruf | Einfach | Schwer | 5. Durchgang | Abruf | Einfach | Schwer |
| Rose - Blume | Salat | _____ | _____ | Frucht - Apfel | Rose | _____ | _____ |
| Salat - Stift | Baby | _____ | _____ | Schule - Bäckerei | Gedränge | _____ | _____ |
| Gehorchen - Meter | Metall | _____ | _____ | Rose - Blume | Baby | _____ | _____ |
| Frucht - Apfel | Schule | _____ | _____ | Salat - Stift | Metall | _____ | _____ |
| Schule - Bäckerei | Rose | _____ | _____ | Metall - Eisen | Gehorchen | _____ | _____ |
| Metall - Eisen | Gedränge | _____ | _____ | Gedränge - Dunkel | Salat | _____ | _____ |
| Gedränge - Dunkel | Frucht | _____ | _____ | Baby - Geschrei | Schule | _____ | _____ |
| Baby - Geschrei | Gehorchen | _____ | _____ | Gehorchen - Meter | Frucht | _____ | _____ |
| Gesamt | | _____ | _____ | Gesamt | | _____ | _____ |
| 3. Durchgang | Abruf | Einfach | Schwer | 6. Durchgang | Abruf | Einfach | Schwer |
| Baby - Geschrei | Gehorchen | _____ | _____ | Metall - Eisen | Baby | _____ | _____ |
| Gedränge - Dunkel | Frucht | _____ | _____ | Rose - Blume | Frucht | _____ | _____ |
| Schule - Bäckerei | Baby | _____ | _____ | Gedränge - Dunkel | Salat | _____ | _____ |
| Rose - Blume | Metall | _____ | _____ | Baby - Geschrei | Rose | _____ | _____ |
| Salat - Stift | Gedränge | _____ | _____ | Gehorchen - Meter | Schule | _____ | _____ |
| Frucht - Apfel | Schule | _____ | _____ | Frucht - Apfel | Gehorchen | _____ | _____ |
| Gehorchen - Meter | Rose | _____ | _____ | Salat - Stift | Gedränge | _____ | _____ |
| Metall - Eisen | Salat | _____ | _____ | Schule - Bäckerei | Metall | _____ | _____ |
| Gesamt | | _____ | _____ | Gesamt | | _____ | _____ |
| Gesamt Durchgänge 1-3 | Max. Einf. = 12 | Max. Schw. = 12 | Max. Gesamt = 24 | | | | |

| VISUELLE WIEDERGABE I Benutzen Sie das Protokollblatt VRI. | | |
|------------------------------------------------------------|-------------------------------------------------------|----------------|
| Händigkeit: _____ Rechts _____ Links | | |
| Aufgabe | Punkte (s. Auswertungsblatt der Visuellen Wiedergabe) | Beobachtungen: |
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | | |
| Max. = 41 Gesamt | | |

Abbildung 13.4.: Testbogen "Verbale Paarerkenntung 1" der WMS-R
(aus: Härting et al. 2000)

13. Verwendete Test- und Untersuchungsbögen

Auswertungsblatt – AKT

Datum:..... Nr.:.....
 Name:..... Vorname:.....
 Alter:..... Diagnose:.....

| | Rohwert | PR | C-Wert |
|----------------------|---------|-------|--------|
| Bearbeitungszeit (T) | | | |
| Richtige (R) | | | |
| Fehler (F) | | _____ | _____ |
| F ₁ | | _____ | _____ |
| F ₂ | | _____ | _____ |
| F ₃ | | _____ | _____ |
| Fehlerprozent (F %) | | | |
| Gesamtmenge (G) | | | |

$$F \% = \frac{F}{R + F} \cdot 100$$

$$G = 35 - F + R$$

Abbildung 13.5.: Auswertungsblatt A-K-T
(aus: Gatterer 1990)

13. Verwendete Test- und Untersuchungsbögen

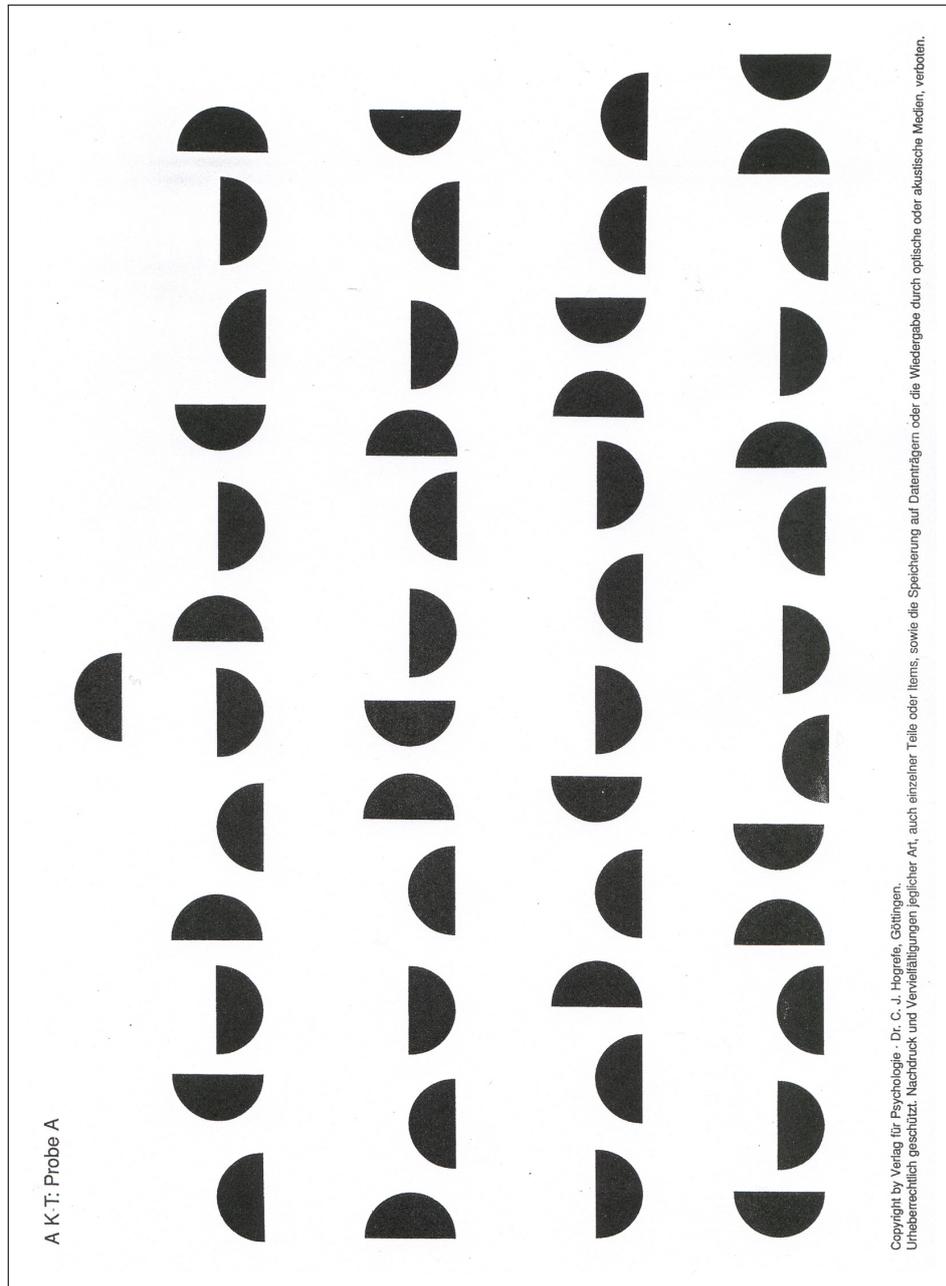


Abbildung 13.6.: Probetestbogen A-K-T
(aus: Gatterer 1990)

13. Verwendete Test- und Untersuchungsbögen

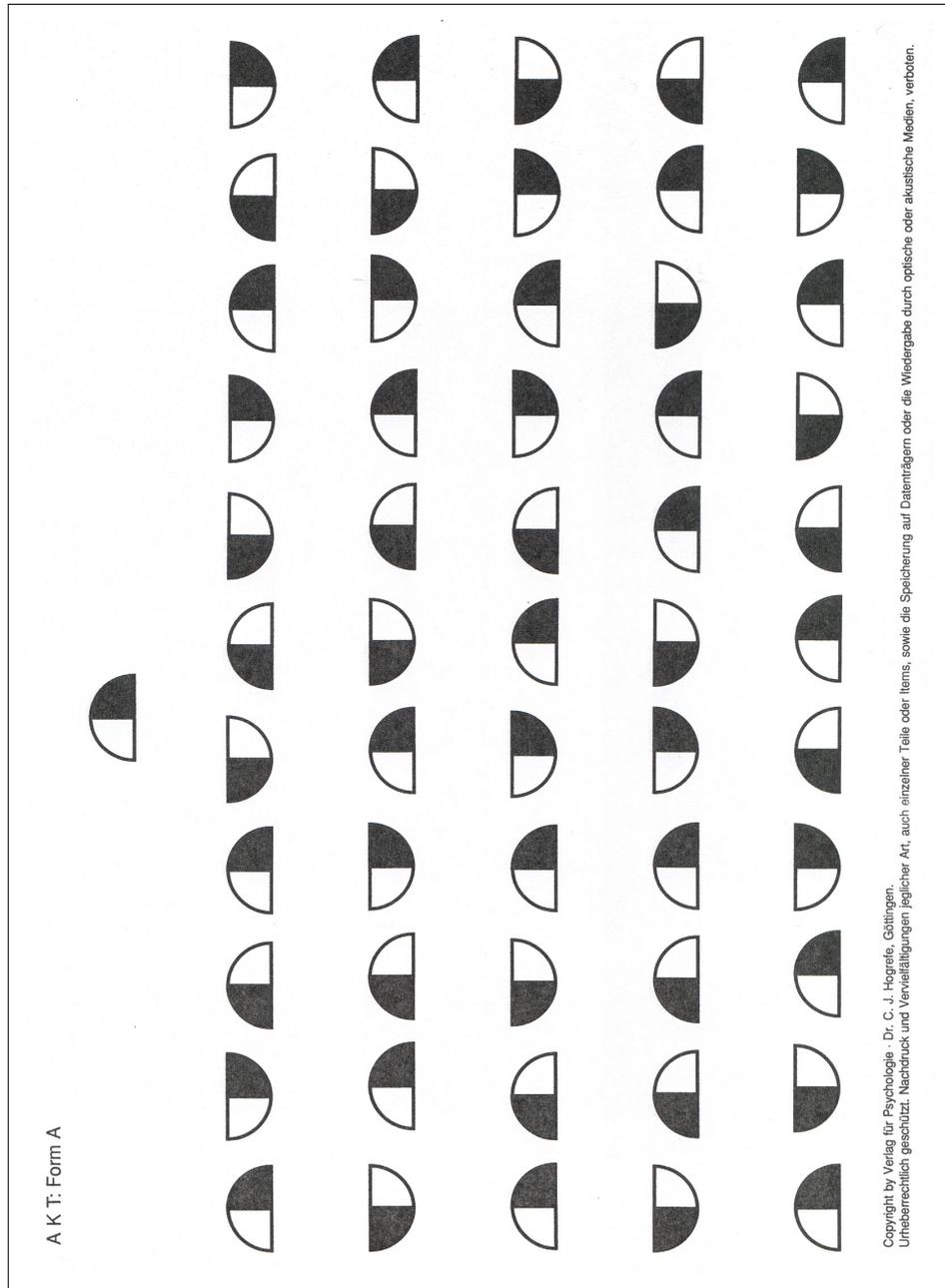


Abbildung 13.7.: Testbogen Form A
(aus: Gatterer 1990)

13. Verwendete Test- und Untersuchungsbögen

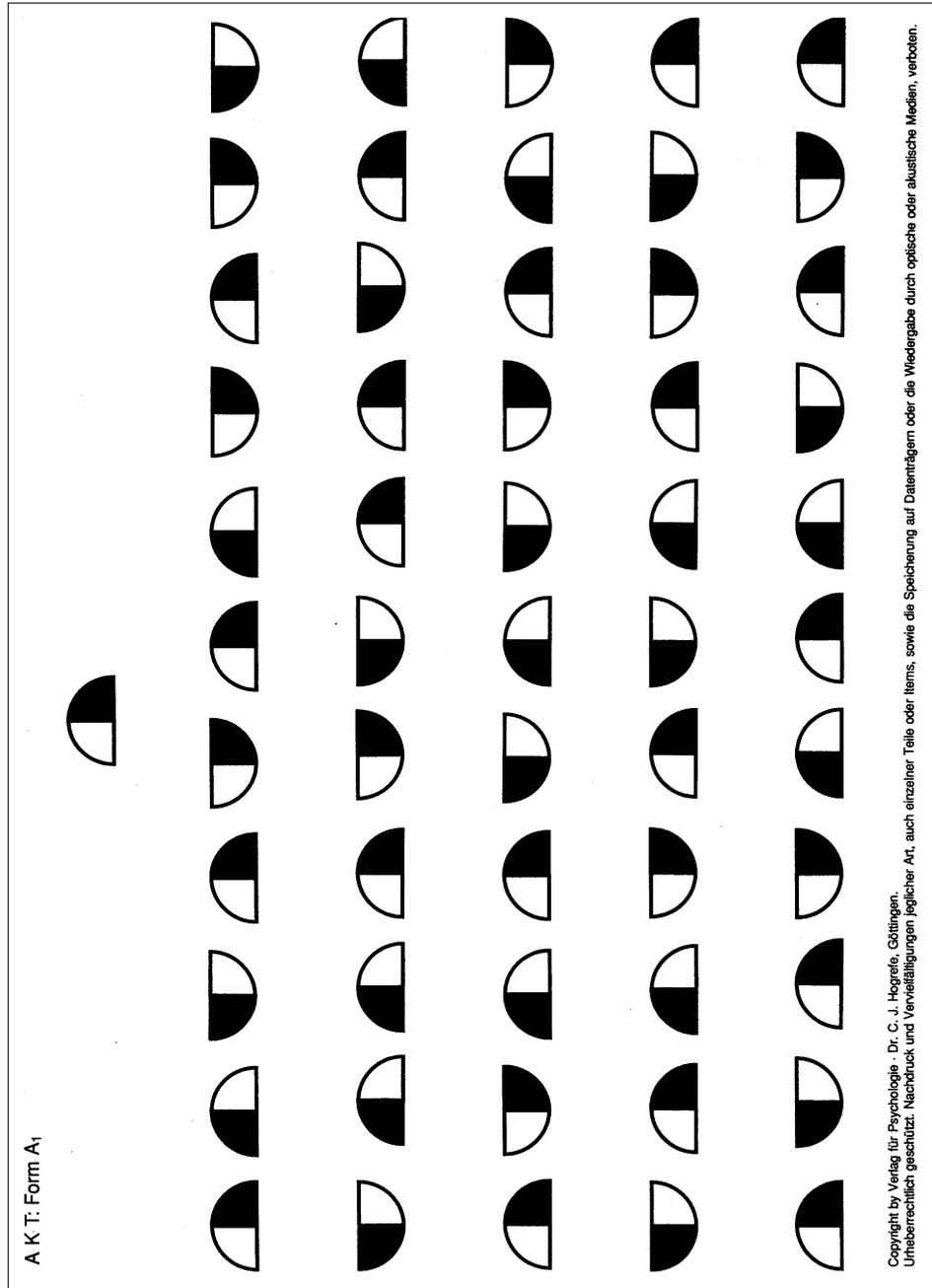


Abbildung 13.8.: Testbogen Form A₁
(aus: Gatterer 1990)

13. Verwendete Test- und Untersuchungsbögen

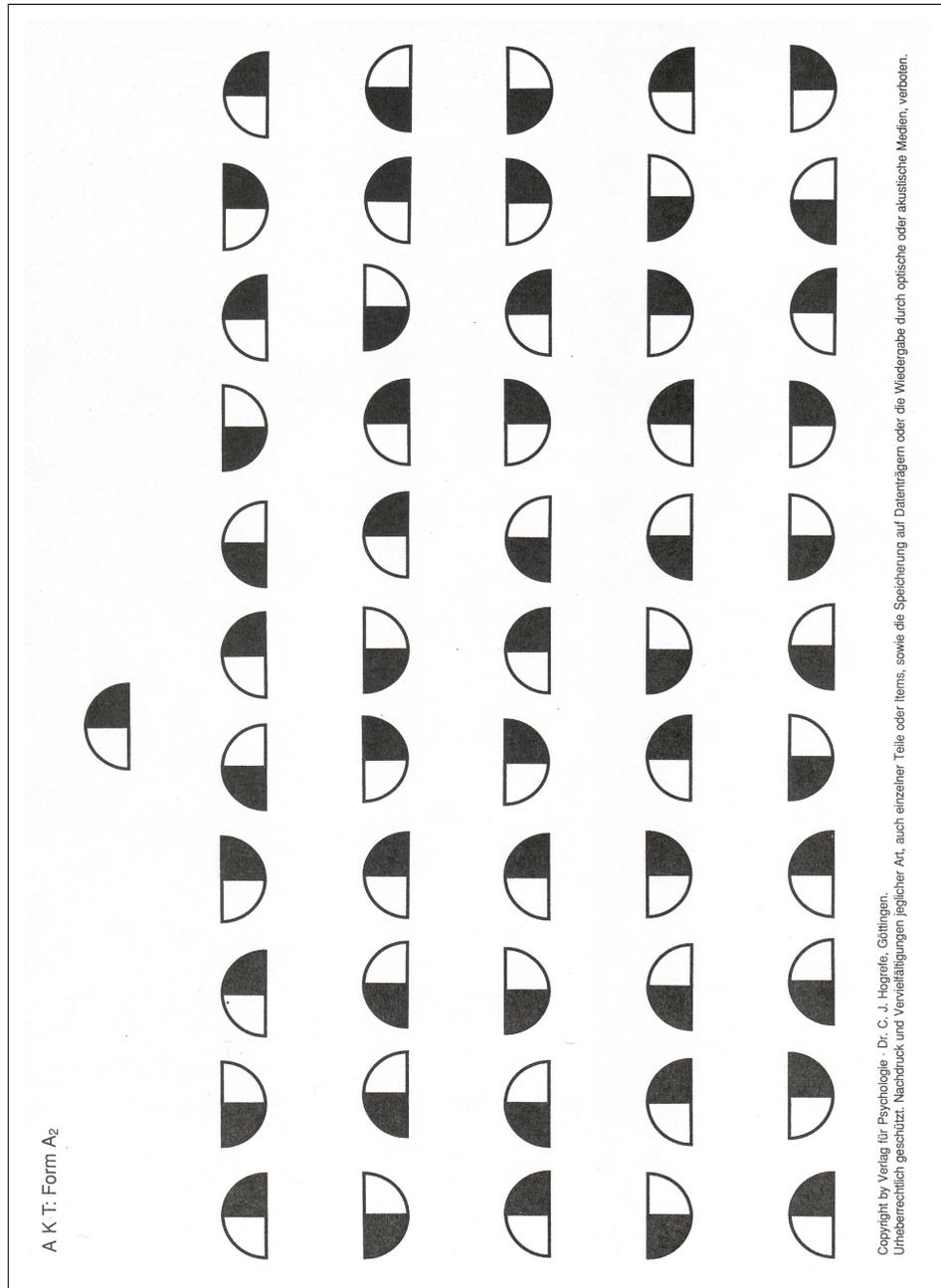


Abbildung 13.9.: Testbogen Form A₂
(aus: Gatterer 1990)

13. Verwendete Test- und Untersuchungsbögen

Fragen zum Befinden

Die folgenden Fragen beziehen sich auf Ihre allgemeine und seelische Verfassung.

Wir bitten Sie, jede Frage zu beantworten und zwar so, wie es für Sie persönlich **in der letzten Woche** am ehesten zutraf.

Machen Sie bitte nur **ein** Kreuz pro Frage und lassen Sie bitte keine Frage aus. Überlegen Sie bitte nicht lange, sondern wählen Sie die Antwort aus, die Ihnen auf Anhieb am zutreffendsten erscheint.

1. Ich kann mich heute noch so freuen wie früher.

- ganz genau so
- nicht so sehr
- nur noch ein wenig
- kaum oder gar nicht

2. Ich kann lachen und die lustige Seite der Dinge sehen.

- ja, soviel wie immer
- nicht mehr ganz so viel
- inzwischen viel weniger
- überhaupt nicht

3. Ich fühle mich glücklich

- meistens
- manchmal
- selten
- überhaupt nicht

Abbildung 13.10.: Testbogen der HADS-D, Subtest Depression (Seite 1)
(Abschrift nach Herrmann-Lingen, Buss & Snaith 1995)

13. Verwendete Test- und Untersuchungsbögen

4. Ich fühle mich in meinen Aktivitäten gebremst.

- überhaupt nicht
- manchmal
- sehr oft
- fast immer

5. Ich habe das Interesse an meiner äußeren Erscheinung verloren.

- ich kümmere mich so viel darum wie immer
- möglicherweise kümmere ich mich zu wenig darum
- ich kümmere mich nicht so sehr darum, wie ich sollte
- ja, stimmt genau

6. Ich blicke mit Freude in die Zukunft.

- ja, sehr
- eher weniger als früher
- viel weniger als früher
- kaum bis gar nicht

7. Ich kann mich an einem guten Buch, einer Radio- oder Fernsehsendung freuen

- oft
- manchmal
- eher selten
- sehr selten

Abbildung 13.11.: Testbogen der HADS-D, Subtest Depression (Seite 2)
(Abschrift nach Herrmann-Lingen, Buss & Snaith 1995)

13. Verwendete Test- und Untersuchungsbögen

| Probandennr. | | | | | |
|--------------|-----------|--------------------------|--------------------------|--------------------------|----------------|
| Nr. | Item | sof. Erk. | verz. Erk. | nicht Erk. | D _x |
| 1 | V1 | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | |
| 2 | V2 | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | |
| 3 | CVC1 | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | |
| 4 | CVC2 | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | |
| 5 | W2 | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | |
| 6 | W1 | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | |
| 7 | S2 | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | |
| 8 | S1 | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | |

Daten erfasst

Abbildung 13.12.: Bogen zur Erfassung der Wiedererkennungsleistung

13. Verwendete Test- und Untersuchungsbögen

Sie haben nun unterschiedliche Stimmuster gehört und sich für ein Muster entschieden, von dem Sie glauben, dass es von Ihrem Ehe-/Lebenspartner stammt
Bitte markieren Sie durch ein Kreuz, wie schwer Ihnen diese Aufgabe gefallen ist (extrem schwer - extrem leicht) und wie sicher Sie sich mit Ihrem Urteil sind (extrem unsicher - extrem sicher).

Aufgabe 1 (V1)

| | | | | | | |
|-----------------|-----------------------------|-----------------------------|----------------------------|----------------------------|----------------------------|---------------|
| extrem schwer | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem leicht |
| extrem unsicher | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem sicher |

Aufgabe 2 (V2)

| | | | | | | |
|-----------------|-----------------------------|-----------------------------|----------------------------|----------------------------|----------------------------|---------------|
| extrem schwer | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem leicht |
| extrem unsicher | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem sicher |

Aufgabe 3 (CVC1)

| | | | | | | |
|-----------------|-----------------------------|-----------------------------|----------------------------|----------------------------|----------------------------|---------------|
| extrem schwer | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem leicht |
| extrem unsicher | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem sicher |

Aufgabe 4 (CVC2)

| | | | | | | |
|-----------------|-----------------------------|-----------------------------|----------------------------|----------------------------|----------------------------|---------------|
| extrem schwer | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem leicht |
| extrem unsicher | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem sicher |

Abbildung 13.13.: Bewertungsbogen (Seite 1)

13. Verwendete Test- und Untersuchungsbögen

Fortsetzung

Aufgabe 5 (W2)

| | | | | | | |
|-----------------|-----------------------------|-----------------------------|----------------------------|----------------------------|----------------------------|---------------|
| extrem schwer | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem leicht |
| extrem unsicher | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem sicher |

Aufgabe 6 (W1)

| | | | | | | |
|-----------------|-----------------------------|-----------------------------|----------------------------|----------------------------|----------------------------|---------------|
| extrem schwer | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem leicht |
| extrem unsicher | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem sicher |

Aufgabe 7 (S2)

| | | | | | | |
|-----------------|-----------------------------|-----------------------------|----------------------------|----------------------------|----------------------------|---------------|
| extrem schwer | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem leicht |
| extrem unsicher | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem sicher |

Aufgabe 8 (S1)

| | | | | | | |
|-----------------|-----------------------------|-----------------------------|----------------------------|----------------------------|----------------------------|---------------|
| extrem schwer | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem leicht |
| extrem unsicher | <input type="checkbox"/> -2 | <input type="checkbox"/> -1 | <input type="checkbox"/> 0 | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | extrem sicher |

Abbildung 13.14.: Bewertungsbogen (Seite 2)

14. Verwendete Hard- und Software

14.1. Verwendete Hardware

| Das Mikrofonmodul | |
|------------------------------------|------------------------------------------|
| Hersteller | Firma Sennheiser, Wedemark (Deutschland) |
| Modell | ME 64 |
| Übertragungsbereich | 40 - 20.000 Hz ($\pm 2,5$ dB) |
| Richtcharakteristik | Niere |
| Freifeldleerlaufübertragungsfaktor | 32 mV/Pa $\pm 2,5$ dB |
| Ersatzgeräuschpegel | 16 dB (nach DIN IEC 651) |
| | 25 dB (nach CCIR 468-3) |
| Grenzschalldruck | 130 dB(A) bei 1 kHz |

Tabelle 14.1.: Technische Daten des verwendeten Mikrofonmoduls

| Der Speiseadapter | |
|-------------------------------|-----------------------------------------------|
| Hersteller | Firma Sennheiser, Wedemark (Deutschland) |
| Modell | K6 |
| Elektrische Impedanz | ca. 200 Ω sym., ca. 100 Ω asym. |
| Empf. Abschlussimpedanz | > 1 k Ω |
| Übertragungsfaktor | 0 db |
| max. Leerlaufausgangsspannung | 2 V (entspricht 6 dBV) |
| Ausgang | symmetrisch transformatorlos |

Tabelle 14.2.: Technische Daten des verwendeten Speiseadapters

14. Verwendete Hard- und Software

| Der Digital Audio Tape-corder (DAT) | |
|-------------------------------------|--------------------------------------------------|
| Hersteller | Firma SONY (Japan) |
| Modell | Walkman TCD-D100 |
| Band | Digitales Audioband |
| Aufnahmezeit | Standardplay: 120 Minuten |
| Abtastfrequenz | 48 kHz |
| Quantisierung | Standardplay: 16 Bit, linear |
| Frequenzgang (Wiedergabe) | Fs 48 kHz 20 – 22.000 Hz ($\pm 1,0$ dB) |
| Signal-Rauschabstand | > 87 dB (1 kHz, IHF-A, LINE IN) |
| Dynamikumfang | > 87 dB (1 kHz, IHF-A, LINE IN) |
| Gesamtklirrgrad | < 0,008% (1 kHz, 22 kHz Tiefpassfilter, LINE IN) |
| Gleichlaufschwankungen | < $\pm 0,001\%$ W. PEAK |
| | Impedanz 4,7 k Ω |
| Eingang | Nenn-Eingangspegel 1,4 mV |
| | Eingangspegel 0,3 mV |

Tabelle 14.3.: Technische Daten des verwendeten DAT-Rekorders

14.2. Verwendete Software

Praat

Bei Praat handelt es sich um eine Software, die von Paul Boersma & David Weenink vom Institute of Phonetic Sciences der Universität Amsterdam, Niederlande, entwickelt wurde und ständig überarbeitet wird. Sie erlaubt die unterschiedlichsten Messungen und Manipulationen sprachlicher und nicht-sprachlicher Signale und stellt dazu zahlreiche unterschiedliche Werkzeuge zur Verfügung. Im Rahmen der vorliegenden Untersuchung wurden die Versionen 4.0 bis 4.4.20 verwendet. Der lizenzfreie Download ist unter dem Link <http://www.praat.org> möglich (letzter Zugriff: 10.05.2006).

Oliomat

Bei dem Oliomat handelt es sich um eine Software zur randomisierten Präsentation akustischer, digitaler Stimuli. Über ein Menüfeld können die gewünschten wav-Dateien ausgewählt und die Pausen zwischen den Stimuli manuell eingestellt werden. Die Software wurde dem Untersucher freundlicherweise von der Firma Thomatechnik, Pfungstadt, kostenfrei zur Verfügung gestellt.

Erklärung

Ich versichere, dass ich die vorliegende Dissertation „**Untersuchung zur Stimmerkennung bei aphasischen und nicht-aphasischen Patienten**“ weder in der gegenwärtigen noch in einer anderen Fassung einer anderen Hochschule als Qualifikationsarbeit zur Erlangung des Doktorgrades oder eines anderen akademischen Grades vorliegt oder Gegenstand eines abgeschlossenen akademischen Qualifikationsverfahrens ist. Ich versichere, dass ich diese Dissertation selbstständig und ohne unerlaubte fremde Hilfe verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel verwendet und die den Quellen wörtlich und inhaltlich entnommenen Stellen kenntlich gemacht habe.

Oliver Kneidl
Bielefeld, den 9. August 2006