

Carl von Ossietzky
Universität Oldenburg

Studiengang Hörtechnik und Audiologie
Master of Science (M.Sc.)

Masterarbeit

Titel:

Perzeptive Beurteilung von Sprache im Störgeräusch

vorgelegt von:
Anne Schlüter (Mat.Nr: 9262490)

Betreuender Gutachter: Frau Prof. Dr. Inga Holube

Zweiter Gutachter: Herr Dr. Thomas Brand

Oldenburg, 21.12.2007

Perzeptive Beurteilung von Sprache im Störgeräusch

Masterarbeit
von
ANNE SCHLÜTER
(Mat.Nr: 9262490)

Ausgabetermin: 05.07.2007
Abgabetermin: 21.12.2007

Betreuer und Erstgutachter: Frau Prof. Dr. Inga Holube
Zweiter Gutachter: Herr Dr. Thomas Brand

Carl von Ossietzky
Universität Oldenburg
D-26111 Oldenburg

Institut für Hörtechnik und Audiologie
Fachhochschule Oldenburg / Ostfriesland / Wilhelmshaven
Fachbereich Bauwesen und Geoinformation (B+G)
Ofener Str. 16
D-26121 Oldenburg

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe.

Oldenburg, den 21.12.2007

Anne Schlüter

Inhaltsverzeichnis

| | |
|--|-----------|
| Abbildungsverzeichnis | 1 |
| Tabellenverzeichnis | 4 |
| 1 Einleitung | 5 |
| 2 Hintergrund | 8 |
| 2.1 Zusammenfassung | 8 |
| 2.2 Einleitung | 9 |
| 2.3 Einkanalige Störgeräuschreduktionen | 9 |
| 2.3.1 Optimalfilter (Wienerfilter) | 10 |
| 2.3.2 Spektrale Subtraktion und ihr Zusammenhang mit dem Optimalfilter | 12 |
| 2.3.3 Realisierung des Optimalfilters und der Spektralen Subtraktion | 13 |
| 2.3.4 Störgeräuschreduktion nach Ephraim und Malah | 14 |
| 2.3.5 Methoden zur Störgeräuschschätzung | 16 |
| 2.4 Perzeptive Maße zur Untersuchung von Sprache im Störgeräusch | 20 |
| 2.4.1 Acceptable Noise Level | 20 |
| 2.4.2 Sprachverständlichkeitstests im Störgeräusch | 23 |
| 2.4.3 Just Follow Conversation Test | 24 |
| 2.4.4 Performance Perceptual Test | 25 |
| 2.4.5 Paarvergleiche | 27 |
| 2.5 Technische Maße zur Untersuchung von Sprache im Störgeräusch | 28 |
| 2.5.1 SNR nach Hagerman und Olofsson | 28 |
| 2.5.2 Modulationstransferfunktion | 29 |
| 3 Untersuchung des Acceptable Noise Level Test | 31 |
| 3.1 Zusammenfassung | 31 |
| 3.2 Einleitung | 32 |
| 3.3 Allgemeine Methodik | 33 |
| 3.3.1 Apparatur | 33 |
| 3.3.2 Versuchspersonen | 33 |
| 3.3.3 Testmaterial | 33 |
| 3.3.4 Durchführung des ANLT | 34 |
| 3.3.5 Durchführung des OLSA | 34 |
| 3.3.6 Durchführung des JFCT | 35 |
| 3.4 Experiment A: Untersuchung der Schrittgröße | 35 |
| 3.5 Experiment B: Untersuchung einer auf- oder absteigenden Instruktionen im ANLT | 39 |
| 3.6 Experiment C: Untersuchung des Startpegels | 42 |
| 3.7 Experiment D: Test-Retest-Reliabilität und Reliabilität über einen kurzen Zeitraum | 44 |

| | | |
|----------|--|------------|
| 3.8 | Experiment E: Einfluss der Methodik bei Wiederholungsmessungen auf den ANL | 52 |
| 3.9 | Experiment F: Einfluss der SRT-Messung auf ANL- und JFC-Werte | 56 |
| 3.10 | Experiment G: Einfluss unterschiedlicher Hintergrundgeräusche | 60 |
| 3.11 | Experiment H: Einfluss des Darbietungspegels bei der Übung des Sprachmaterials | 64 |
| 3.12 | Schlussfolgerungen | 66 |
| 4 | Störgeräuschreduktionen, perzeptive Maße | 68 |
| 4.1 | Zusammenfassung | 68 |
| 4.2 | Einleitung | 69 |
| 4.3 | Methode | 71 |
| 4.3.1 | Versuchspersonen | 71 |
| 4.3.2 | Aufbau und Testmaterial | 72 |
| 4.3.3 | Einkanalige Störgeräuschreduktionen | 73 |
| 4.3.4 | Messverfahren | 75 |
| 4.4 | Ergebnisse | 77 |
| 4.5 | Diskussion | 86 |
| 4.6 | Schlussfolgerungen | 89 |
| 5 | Störgeräuschreduktionen, technischen Maße | 90 |
| 5.1 | Zusammenfassung | 90 |
| 5.2 | Einleitung | 91 |
| 5.3 | Methode | 92 |
| 5.3.1 | Versuchspersonen | 92 |
| 5.3.2 | Aufbau und Testmaterial | 92 |
| 5.3.3 | Einkanalige Störgeräuschreduktionen | 94 |
| 5.3.4 | Perzeptives Messverfahren: ANLT | 95 |
| 5.3.5 | Technische Messverfahren | 97 |
| 5.4 | Ergebnisse | 98 |
| 5.5 | Diskussion | 104 |
| 5.6 | Schlussfolgerungen | 105 |
| 6 | Zusammenfassung und Ausblick | 107 |
| A | Messaufbau, Freifeldentzerrung, Kalibrierung | 110 |
| A.1 | Messaufbau | 110 |
| A.2 | Ermittlung der Freifeldentzerrung | 111 |
| A.3 | Kalibrierung | 111 |
| B | Instruktionen | 112 |
| B.1 | Instruktion des ANLT | 112 |
| B.2 | Verwendete Instruktionen bei der Untersuchung ihres Einflusses | 114 |
| B.2.1 | Instruktion: Aufsteigend | 114 |
| B.2.2 | Instruktion: Absteigend | 115 |
| B.3 | Instruktion der MCL-Messung | 117 |
| B.4 | Instruktion der BNL-Messung | 118 |
| B.5 | Instruktion des Just Follow Conversation Tests | 119 |
| B.6 | Instruktion des Oldenburger Satztests | 120 |
| B.7 | Instruktion des Paarvergleichs | 120 |
| C | Hörverluste der Versuchspersonen | 122 |
| C.1 | Untersuchung der Schrittgröße | 122 |
| C.2 | Untersuchung einer auf- und absteigenden Instruktion | 123 |
| C.3 | Untersuchung des Startpegels | 123 |

| | | |
|----------|---|------------|
| C.4 | Untersuchung der Test-Retest-Reliabilität und Reliabilität über einen kurzen Zeitraum | 124 |
| C.5 | Einfluss der Methodik bei Wiederholungsmessungen auf den ANL | 124 |
| C.6 | Einfluss einer SRT-Messung auf ANL- und JFC-Werte | 125 |
| C.7 | Einfluss unterschiedlicher Hintergrundgeräusche | 125 |
| C.8 | Einfluss des Darbietungspegels bei der Übung des Sprachmaterials | 126 |
| D | Untersuchung des ANLT | 127 |
| D.1 | Experiment A: Untersuchung der Schrittgröße | 127 |
| D.1.1 | Signifikanzprüfung der Messwerte im ANLT | 127 |
| D.1.2 | ANL-Werte gemessen mit unterschiedlichen Schrittgrößen | 128 |
| D.2 | Experiment B: Untersuchung einer auf- oder absteigenden Instruktion | 129 |
| D.2.1 | Messwerte des ANLT, gemessen mit auf- und absteigender Instruktion | 129 |
| D.2.2 | Signifikanzprüfung der Messwerte im ANLT | 130 |
| D.3 | Experiment C: Untersuchung des Startpegels | 130 |
| D.4 | Experiment D: Test-Retest-Reliabilität und Reliabilität über einen kurzen Zeitraum | 131 |
| D.4.1 | Gegenüberstellung der Messwerte des ANLT des Termin A und B | 131 |
| D.4.2 | Inter- und intraindividuelle Streuung des ANL | 132 |
| D.4.3 | Signifikanzprüfung der kumulierten ANL-Werte | 133 |
| D.4.4 | Berücksichtigung der interindividuellen Streuung im kumulierten ANL | 134 |
| D.4.5 | Gegenüberstellung der JFC-Werte der Termine A und B | 135 |
| D.4.6 | Inter- und intraindividuelle Streuung der JFC-Werte | 135 |
| D.4.7 | Signifikanzprüfung der kumulierten JFC-Werte | 136 |
| D.4.8 | Berücksichtigung der interindividuellen Streuung im kumulierten JFC-Wert | 137 |
| D.4.9 | Abhängigkeit von ANLT und JFCT | 137 |
| D.5 | Experiment E: Einfluss der Methodik bei Wiederholungsmessungen auf den ANL | 139 |
| D.5.1 | Gegenüberstellung der Messwerte im ANLT | 139 |
| D.5.2 | Inter- und intraindividuelle Streuung des ANL | 140 |
| D.5.3 | Signifikanzprüfung der kumulierten ANL-Werte | 140 |
| D.5.4 | Berücksichtigung der interindividuellen Streuung im kumulierten ANL | 141 |
| D.6 | Experiment G: Einfluss unterschiedlicher Hintergrundgeräusche | 141 |
| D.7 | Experiment H: Einfluss des Darbietungspegels bei der Übung des Sprachmaterials | 142 |
| E | Störgeräuschreduktionen, perzeptive Maße | 143 |
| E.1 | ANL-Werte ermittelt mit Störgeräuschreduktion | 143 |
| E.2 | Signifikanzprüfung der perzeptiven Maße | 144 |
| E.3 | Paarvergleich | 145 |
| E.3.1 | Ergebnisse der Paarvergleiche bei unterschiedlichen SNR-Werten | 145 |
| E.3.2 | Prüfung der Konsistenz und Akkordanz im Paarvergleich | 148 |
| E.4 | Gegenüberstellung der perzeptiven Maße und der tatsächlichen SNR-Verbesserung | 149 |
| E.5 | Vergleich der ANL-Werte mit den anderen perzeptiven Maßen | 155 |
| F | Störgeräuschreduktionen, technische Maße | 159 |
| | Literaturverzeichnis | 162 |

Abbildungsverzeichnis

| | | |
|------|---|----|
| 2.1 | Signalverarbeitung in einer Störgeräuschreduktion | 10 |
| 2.2 | Design des ANL | 21 |
| 2.3 | Design des PPT | 26 |
| 3.1 | ANLT-Messwerte bei unterschiedlichen Schrittgrößen | 37 |
| 3.2 | Differenz zwischen maximalem und minimalem ANL-Wert bei unterschiedlichen Schrittgrößen | 37 |
| 3.3 | Anzahl der Pegelveränderungen bei unterschiedlichen Schrittgrößen | 38 |
| 3.4 | ANLT-Werte bei auf- oder absteigender Instruktion | 41 |
| 3.5 | ANLT-Werte bei unterschiedlichem Startpegel | 43 |
| 3.6 | MCL-, BNL- und ANL-Werte gemessen mit zwölf Wiederholungen | 46 |
| 3.7 | Kumulierter ANL | 47 |
| 3.8 | JFC-Werte, gemessen mit zwölf Wiederholungen | 48 |
| 3.9 | Kumulierter JFC-Wert | 49 |
| 3.10 | Mediane der ANL-Werte (Messreihenfolge: ANLT-JFCT, JFCT-ANLT) | 50 |
| 3.11 | Mediane der JFC-Werte (Messreihenfolge: ANLT-JFCT, JFCT-ANLT) | 51 |
| 3.12 | MCL-, BNL- und ANL-Werte, gemessen mit zwölf Wiederholungen (konstanter MCL-Wert) | 54 |
| 3.13 | Kumulierter ANL-Wert (konstanter MCL-Wert) | 55 |
| 3.14 | Versuchsdesign im Experiment „Einfluss der SRT-Messung“ | 58 |
| 3.15 | JFC-, SRT- und ANL-Werte im Experiment „Einfluss der SRT-Messung“ | 59 |
| 3.16 | Spektren der Hintergrundgeräusche | 62 |
| 3.17 | ANL-Werte gemessen mit unterschiedlichen Hintergrundgeräuschen | 63 |
| 3.18 | ANLT-Werte in der Untersuchung des Einflusses des Darbietungspegels in der Übung | 66 |
| 4.1 | Hörverlust der schwerhörigen Versuchspersonen | 72 |
| 4.2 | Zur Analyse der SNR-Verbesserung genutzte Signale | 74 |
| 4.3 | SNR-Verbesserung durch die Störgeräuschreduktionen | 74 |
| 4.4 | ANL-Werte mit Störgeräuschreduktion, Normalhörende und Schwerhörige | 78 |
| 4.5 | Δ ANL-Werte mit Störgeräuschreduktion, Normalhörende und Schwerhörige | 79 |
| 4.6 | Δ JFC-Werte bei unterschiedlichen Störgeräuschreduktion, Normalhörende und Schwerhörige | 80 |
| 4.7 | Δ SRT-Werte bei unterschiedlichen Störgeräuschreduktionen, Normalhörende und Schwerhörige | 81 |
| 4.8 | Δ PPDIS-Werte bei unterschiedlichen Störgeräuschreduktionen, Normalhörende und Schwerhörige | 82 |
| 4.9 | Gesamtergebnis der Paarvergleiche mit unterschiedlichen Störgeräuschreduktionen, Normalhörende und Schwerhörige | 83 |
| 4.10 | Präferenz gegenüber <i>NoAlgo</i> mit unterschiedlichen Störgeräuschreduktionen, Normalhörende und Schwerhörige | 83 |
| 4.11 | Vergleich perzeptive Maße und tatsächliche SNR-Verbesserung | 85 |

| | | |
|------|---|-----|
| 5.1 | Hörverlust der schwerhörigen Versuchspersonen | 93 |
| 5.2 | Zur Analyse der tatsächlichen SNR-Verbesserung genutzte Signale | 94 |
| 5.3 | Tatsächliche SNR-Verbesserung durch die Störgeräuschreduktionen | 95 |
| 5.4 | ANL-Werte mit Störgeräuschreduktion, Normalhörende und Schwerhörige | 99 |
| 5.5 | Δ ANL-Werte mit Störgeräuschreduktion, Normalhörende und Schwerhörige | 100 |
| 5.6 | Δ SNR und MTF der unterschiedlichen Störgeräuschreduktionen | 101 |
| 5.7 | Δ SNR-Werte, normierte logarithmierte MTF und tatsächliche SNR-Verbesserung der unterschiedlichen Störgeräuschreduktionen | 102 |
| A.1 | Messaufbau | 110 |
| C.1 | Hörvermögen in der Untersuchung der Schrittgröße | 122 |
| C.2 | Hörvermögen in der Untersuchung der Orientierung | 123 |
| C.3 | Hörvermögen in der Untersuchung der Startpegel | 123 |
| C.4 | Hörvermögen in der Untersuchung der Test-Retest-Reliabilität | 124 |
| C.5 | Hörvermögen in der Untersuchung der Methodik bei Wiederholungsmessungen auf den ANL | 124 |
| C.6 | Hörvermögen in der Untersuchung der SRT-Messung | 125 |
| C.7 | Hörvermögen in der Untersuchung unterschiedlicher Hintergrundgeräusche | 125 |
| C.8 | Hörvermögen in der Untersuchung des Darbietungspegels bei der Übung des Sprachmaterials | 126 |
| D.1 | ANL-Werte jeder Versuchsperson bei unterschiedlicher Schrittgröße | 128 |
| D.2 | ANLT-Werte bei auf- oder absteigender Instruktion (Messreihenfolge) | 129 |
| D.3 | Scatterplot der MCL-, BNL- und ANL-Werte (Termin A und B) | 131 |
| D.4 | Inter- und intraindividuelle Streuung des ANL | 132 |
| D.5 | Differenzieller und differenzieller kumulierter ANL | 134 |
| D.6 | Scatterplot der JFC-Werte (Termin A und B) | 135 |
| D.7 | Inter- und intraindividuelle Streuung der JFC-Werte | 135 |
| D.8 | Differenzieller und differenzieller kumulierter JFC-Wert | 137 |
| D.9 | ANL-Werte gemessen mit den Reihenfolgen ANLT-JFCT und JFCT-ANLT | 138 |
| D.10 | JFC-Werte gemessen mit den Reihenfolgen ANLT-JFCT und JFCT-ANLT | 138 |
| D.11 | Scatterplot der MCL-, BNL- und ANL-Werte (Termin A und B, konstanter MCL-Wert) | 139 |
| D.12 | Inter- und intraindividuelle Streuung (konstanter MCL-Wert) | 140 |
| D.13 | Normierter und differenziellerer kumulierter ANL (konstanter MCL-Wert) | 141 |
| E.1 | Ergebnisse des Paarvergleichs bei den SNR-Werten -3, 7,5, 12,5, 19 dB SNR mit unterschiedliche Störgeräuschreduktionen, Normalhörende und Schwerhörige | 146 |
| E.2 | Ergebnisse des Paarvergleichs bei den SNRs: ANL NoAlgo, ANL Real6dB, ANL Real8dB, ANL Optimal mit unterschiedlichen Störgeräuschreduktionen, Normalhörende und Schwerhörige | 147 |
| E.3 | ANL-Werte und SNR-Verbesserung bei unterschiedlichen Störgeräuschreduktionen, Normalhörende und Schwerhörige | 151 |
| E.4 | JFC-Werte und SNR-Verbesserung bei unterschiedlichen Störgeräuschreduktionen, Normalhörende und Schwerhörige | 152 |
| E.5 | SRT-Werte und SNR-Verbesserung bei unterschiedlichen Störgeräuschreduktionen, Normalhörende und Schwerhörige | 153 |
| E.6 | Paarvergleich und SNR-Verbesserung der Störgeräuschreduktionen, Normalhörende und Schwerhörige | 154 |
| E.7 | Scatterplots ANL- und JFC-Werte, Normalhörende und Schwerhörige | 155 |
| E.8 | Scatterplots ANL- und SRT-Werte, Normalhörende und Schwerhörige | 156 |
| E.9 | Scatterplots ANL- und PPDIS-Werte, Normalhörende und Schwerhörige | 157 |

| | | |
|------|---|-----|
| E.10 | Scatterplots ANL-Werte und Werte des Paarvergleichs, Normalhörende und Schwerhörige | 158 |
| F.1 | Δ SNR-Werte, SNR-Verbesserung und Δ ANL-Werte der unterschiedlichen Störgeräuschreduktionen, Normalhörende und Schwerhörige | 160 |
| F.2 | Normierte und logarithmierte MTF, SNR-Verbesserung und Δ ANL-Werte der unterschiedlichen Störgeräuschreduktionen, Normalhörende und Schwerhörige | 161 |

Tabellenverzeichnis

| | | |
|-----|--|-----|
| 2.1 | Design des OLSA | 24 |
| 2.2 | Beispiel einer Dominanzmatrix aus einem Paarvergleich | 28 |
| 3.1 | Signifikanzprüfung des SRT-Einfluss | 59 |
| 4.1 | ANL-Werte der Normal- und Schwerhörigen | 79 |
| 4.2 | Mediane PPDIS | 81 |
| 4.3 | Spearmanischer Korrelationskoeffizient zwischen Ergebnisse des ANLT und den anderen perceptiven Maßen | 86 |
| 5.1 | Spearmanischer Korrelationskoeffizient zwischen Δ SNR und SNR_{OUT-IN} | 103 |
| 5.2 | Spearmanischer Korrelationskoeffizient zwischen Δ ANL und Δ SNR nach Hagerman und Olofsson | 103 |
| D.1 | Signifikanzprüfung: Schrittgröße | 127 |
| D.2 | Signifikanzprüfung: auf-/absteigende Instruktion | 130 |
| D.3 | Signifikanzprüfung: Startpegel | 130 |
| D.4 | Signifikanzprüfung der kumulierten ANL-Werte des Kap. 3.7 | 133 |
| D.5 | Signifikanzprüfung: kumulierter JFC-Wert | 136 |
| D.6 | Signifikanzprüfung der kumulierten ANL-Werte (konstanter MCL-Wert) | 140 |
| D.7 | Signifikanzprüfung der ANL-Werte unterschiedlicher Hintergrundgeräusche | 141 |
| D.8 | Signifikanzprüfung der Messwerte des ANLT bei der Prüfung des Ein- flusses des Darbietungspegels der Sprache in der Übung | 142 |
| E.1 | ANL-Werte Normalhörende | 143 |
| E.2 | ANL-Werte Schwerhörige | 144 |
| E.3 | Signifikanzprüfung perceptiv Maße | 144 |
| E.4 | Konsistenz Normalhörende | 148 |
| E.5 | Konsistenz Schwerhörige | 148 |
| E.6 | Akkordanz Normalhörende | 149 |
| E.7 | Akkordanz Schwerhörige | 149 |
| E.8 | Spearmanischer Korrelationskoeffizient zwischen den perceptiven Maßen und SNR_{OUT-IN} | 150 |

Kapitel 1

Einleitung

Ziel einkanaliger Störgeräuschreduktionen ist eine Verbesserung der Sprachverständlichkeit und/oder der Klangqualität durch Absenkung des Störsignals. Daher werden sie oft bei der Verarbeitung von Sprache im Hintergrundgeräusch eingesetzt und z. B. in Hörgeräten verwendet (Bentler und Chiou, 2006). Einkanalige Störgeräuschreduktionen bestehen in der Regel aus einer Störgeräuschschätzung und einer Verstärkungsregel. Die Störgeräuschschätzung stellt im vorliegenden gestörten Sprachsignal den Hintergrundgeräuschanteil fest. Methoden der Schätzung sind z. B. Minimum Statistics (Martin, 1993, 1994, 2001, 2003), Amplitudenmodulationsspektrogramme (Tchorz und Kollmeier, 2003), Voice Activity Detection und Periodogramme (Appiah *et al.*, 2005) oder Minima Controlled Recursive Averaging (Cohen und Berdugo, 2002). In einem zweiten Schritt wird nach der Störgeräuschschätzung die frequenzabhängige Verstärkung berechnet. Berechnungsformeln hierzu sind z. B. die Spektrale Subtraktion, der Optimalfilter (Vary *et al.*, 1998) oder die Störgeräuschreduktion nach Ephraim und Malah (1984, 1985). Diese Verstärkungsregeln reduzieren den Pegel in verschiedenen Frequenzbändern, die vom Hintergrundgeräusch dominiert werden. Der Nutzen einkanaliger Störgeräuschreduktionen kann mit perzeptiven und technischen Maßen untersucht werden, wobei unterschiedliche Problematiken auftreten.

Eine naheliegende perzeptive Methode zur Untersuchung der Verbesserung der Sprachverständlichkeit durch Störgeräuschreduktionen ist ein Sprachverständlichkeitstest im Störgeräusch. Ziel dieses Tests ist z. B. die Ermittlung eines Signal-Rausch-Verhältnisses (engl. *Signal-to-Noise Ratio*, *SNR*), bei dem 50% der Sprache verstanden wird. Der erreichte SNR wird Sprachverständlichkeitsschwelle genannt. Unterschiedliche Studien, in denen das Sprachverstehen mit Sprachverständlichkeitstests geprüft wurde, zeigten bei der Untersuchung unterschiedlicher einkanaliger Störgeräuschreduktionen keine (Eberling *et al.*, 1993; Boymans *et al.*, 1999; Marzinzik, 2000; Walden *et al.*, 2000; Alcántara *et al.*, 2003; Ricketts und Hornsby, 2005; Mueller *et al.*, 2006) oder nur eine geringe Verbesserung der Sprachverständlichkeit (Bray und Nilsson, 2001; Lurquin *et al.*, 2001; Bray *et al.*, 2002). Erklärt wurden diese Ergebnisse u. a. durch die Untersuchungsmethode, den Sprachverständlichkeitstest. In ihm wurden Sprachverständlichkeitsschwellen gemessen, die vor allem für Normalhörende bei negativen SNR-Werten liegen (Kollmeier und Wesselkamp, 1997; Wagener *et al.*, 1999a,b,c). D. h., das Hintergrundgeräusch

hat einen höheren Pegel als die Sprache. Einkanalige Störgeräuschreduktionen arbeiten jedoch am effektivsten bei positiven SNR-Werten (Marzinzik, 2000).

Durch Störgeräuschreduktionen könnte auch die subjektiv wahrgenommene Sprachverständlichkeit beeinflusst werden. Sie kann z. B. mit dem Just Follow Conversation Test (JFCT) dokumentiert werden (Larsby und Arlinger, 1994; Borg *et al.*, 1998; Magnusson *et al.*, 2001; Neumann und Jensen, 2001). In ihm wird der Versuchsperson ein Sprachsignal zusammen mit einem Hintergrundgeräusch dargeboten. Sie hat die Aufgabe, den Pegel der Sprache zu verändern, bis sie die Sprache gerade eben noch versteht. Dieser Vorgang unterliegt den subjektiven Maßstäben der Versuchsperson. Dokumentiert wird der SNR der Einstellung. Verschiedene Studien zeigen jedoch, dass auch mit diesem Test negative SNR-Werte erreicht werden können (Larsby und Arlinger, 1994; Neumann und Jensen, 2001; Latzel *et al.*, 2003). Die einkanaligen Störgeräuschreduktionen können daher auch bei diesem Verfahren nicht in einem optimalen Bereich arbeiten.

Eine Messgröße, die die objektiv gemessene und die individuell wahrgenommene Sprachverständlichkeit miteinander vergleicht, ist die Perceptual Performance Discrepancy (PPDIS). Dies ist nach Saunders und Cienkowski (2002) die Differenz zwischen einer subjektiv wahrgenommenen und der objektiv gemessenen Sprachverständlichkeitsschwelle. Versuchspersonen mit einer positiven PPDIS überschätzen ihre Hörfähigkeit, und Versuchspersonen mit einer negativen PPDIS unterschätzen sie.

Zur Untersuchung der Veränderung der Klangqualität durch einkanalige Störgeräuschreduktionen werden häufig Paarvergleiche durchgeführt (Boymans und Dreschler, 2000; Ricketts und Hornsby, 2005). In ihnen werden Signalkaare dargeboten, und die Versuchsperson hat z. B. das Signal mit der besseren Gesamtqualität zu wählen. Dieses Verfahren ist auch für geringe wahrnehmbare Unterschiede sensitiv. Da beim Paarvergleich jedes Signal mit jedem anderen Signal verglichen werden muss, um eine Rangordnung der Signale zu erhalten, steigt die Anzahl der Vergleiche mit der Anzahl der verwendeten Signale schnell an.

Neben der Klangqualität und der Sprachverständlichkeit nehmen einkanalige Störgeräuschreduktionen Einfluss auf die Akzeptanz von Hintergrundgeräuschen (Mueller *et al.*, 2006). Zur Untersuchung dieses Effektes kann der Acceptable Noise Level Test (ANLT) verwendet werden (Nabelek *et al.*, 1991, 2006). In diesem Verfahren wird der akzeptierbare SNR für Sprache im Hintergrundgeräusch ermittelt, der Acceptable Noise Level (ANL) genannt wird. Er ist definiert als Differenz zwischen dem Most Comfortable Level (MCL) der Sprache und dem höchsten Pegel des Hintergrundgeräusches, den die Versuchsperson über längere Zeit beim Hören von Sprache akzeptieren würde. Versuchspersonen mit einem niedrigen ANL tolerieren höhere Hintergrundgeräuschpegel als Versuchspersonen mit einem hohen ANL.

Zur technischen Untersuchung von Störgeräuschreduktionen eignen sich Modelle und Maße, die die SNR-Veränderung untersuchen. Dazu wurden z. B. der Δ SNR nach Hagerman und Olofsson (2004) oder die Modulationstransferfunktion (MTF) nach Drullmann *et al.* (1994) vorgeschlagen. Der Δ SNR ist ein Maß für die Veränderung des SNR durch ein Übertragungssystem. Houtgast und Steeneken (1985) stellten die MTF als ein

Maß vor, mit dem die Veränderung der Modulationstiefe durch ein Übertragungssystem beschrieben werden kann.

Diese Arbeit untersucht den ANLT zur Dokumentation des Nutzens von einkanaligen Störgeräuschreduktionen. Seine Methodik wird in verschiedenen Experimenten geprüft. Anschließend wird er auf die Untersuchung von Störgeräuschreduktionen angewendet und seine Ergebnisse denen des OLSA, des JFCT, der PPDIS und des Paarvergleichs gegenübergestellt. Untersucht wird, ob der ANLT zur Entwicklung von einkanaligen Störgeräuschreduktionen verwendet werden kann und z. B. eindeutigere Ergebnisse als ein Sprachverständlichkeitstest erreicht oder weniger aufwändig ist als ein Paarvergleich. Außerdem wird der ANLT mit den technischen Maßen Δ SNR und MTF verglichen.

Kap. 2 stellt die Hintergründe und Zusammenhänge verschiedener Methoden und Strategien der Störgeräuschreduktion, perceptiver und technischer Maße zusammen. Es ist als Hilfestellung für den interessierten Leser gedacht, der zu Kap. 3, 4 und 5 mehr Hintergrundinformation wünscht.

Kap. 3 untersucht das perzeptive Verfahren des ANLT. In unterschiedlichen Experimenten werden die Einflüsse der Schrittgröße, der Reihenfolge der Instruktion, des Startpegels, unterschiedlicher Hintergrundgeräusche und der Übung des Sprachmaterials auf den ANL untersucht. Außerdem werden die Test-Retest-Reliabilität des ANLT und seine Reliabilität über einen kurzen Zeitraum, die Methodik bei Messwiederholungen und der Einfluss von SRT-Ermittlungen mit dem Oldenburger Satztest und dem Just Follow Conversation Test auf dem ANL geprüft.

Kap. 4 prüft die perceptiven Messverfahren ANLT, JFCT, OLSA, PPDIS und Paarvergleich zur Dokumentation des Nutzens von einkanaligen Störgeräuschreduktionen. Der Schwerpunkt wird dabei auf den ANLT gelegt. Seine Ergebnisse werden verglichen mit denen der anderen Verfahren. Geprüft wird ein Algorithmus, der ein a priori Wissen des Hintergrundgeräusches nutzt. Außerdem werden zwei realistische Algorithmen untersucht. Sie schätzen aus dem Gesamtsignal das Rauschen.

Kap. 5 vergleicht die Ergebnisse des perceptiven Messverfahrens ANLT bei der Untersuchung unterschiedlicher Störgeräuschreduktionen (s. Kap. 4) mit den technischen Maßen Modulationstransferfunktion und Signal-Rausch-Abstand nach Hagerman und Olofsson.

Die Kap. 3 - 5 sollen veröffentlicht werden und sind deshalb in der Form von wissenschaftlichen Artikeln geschrieben.

Kapitel 2

Hintergrund

2.1 Zusammenfassung

Ziel einer Störgeräuschreduktion ist die Verbesserung des Signal-Rausch-Verhältnisses (engl. Signal-to-Noise Ratio, SNR) zwischen Nutz- und Störsignal und damit eine Verbesserung der Sprachverständlichkeit und der Klangqualität. Eingangssignal einer einkanaligen Störgeräuschreduktion ist das gestörte Nutzsinal, das mit Hilfe mehrerer Verarbeitungsschritte entstört wird. Dazu ist zuerst die Leistung des Störsignals zu schätzen. Die geschätzte Rauschleistung wird dann in einer Verstärkungsregel verwendet. Der Nutzen der Störgeräuschreduktionen kann mit unterschiedlichen perzeptiven und technischen Tests untersucht werden.

Eine naheliegende perzeptive Untersuchungsform sind Sprachverständlichkeitstests im Störgeräusch. Mit ihnen wird ein SNR ermittelt, bei dem 50% des Sprachmaterials verstanden wird. Eine andere Form des Sprachverständlichkeitstests im Störgeräusch ist der Just Follow Conversation Test (JFCT), der eine subjektive Sprachverständlichkeitsschwelle ermittelt. Die Versuchsperson stellt selbst den Pegel der Sprache oder des Rauschens ein, bei dem sie die Sprache gerade eben verstehen kann. Ein Vergleich der Werte eines Sprachverständlichkeitstest und der subjektiven Einstellung der Sprachverständlichkeitsschwelle, wie z. B. bei JFCT, führt der Performance Perceptual Test durch. Eine andere Möglichkeit, Sprache im Störgeräusch zu untersuchen, ist der Acceptable Noise Level Test (ANLT). In ihm wird der Acceptable Noise Level (ANL) ermittelt, der definiert ist als Differenz zwischen dem angenehmen Pegel der Sprache und dem maximal tolerierbaren Pegel des Hintergrundgeräusches. Eine Möglichkeit, die Signalqualität zu untersuchen ist der Paarvergleich. Dazu werden die Stimuli in Paaren dargeboten und von der Versuchsperson ein Stimulus gewählt, der eine deutlichere Ausprägung eines Merkmals (z. B. die Signal- oder Sprachqualität) besitzt. Das Ergebnis ist eine Rangfolge der Stimuli. Unterschiedliche technische Messungen können ebenso verwendet werden, um den Nutzen von Störgeräuschreduktionen zu dokumentieren. Der Δ SNR nach Hagermann und Olofsson untersucht die SNR-Veränderung durch eine Störgeräuschreduktion. Mit der Modulationstransferfunktion können die Veränderungen der Modulationen durch ein Übertragungssystem gemessen werden.

2.2 Einleitung

Einkanalige Störgeräuschreduktionen werden in unterschiedlichen Situationen eingesetzt, so z.B. in Freisprechanlagen von Telefonen oder in Hörgeräten (Bentler und Chiou, 2006). Ihre Aufgabe ist eine Verbesserung der Sprachverständlichkeit und Klangqualität durch Filterung des Störsignals. Einkanalige Störgeräuschreduktionen schätzen das Rauschsignal und entfernen es dann durch Filterung aus dem Gesamtsignal (Vary *et al.*, 1998). Zur Dokumentation des Nutzens dieser Algorithmen können sowohl perzeptive als auch technische Maße verwendet werden.

Perzeptive Maße werden z.B. ermittelt im Sprachverständlichkeitstests im Störgeräusch (Wagener *et al.*, 1999c,a,b; Wagener, 2003), im Just Follow Conversation Test (Larsby und Arlinger, 1994; Borg *et al.*, 1998; Magnusson *et al.*, 2001; Neumann und Jensen, 2001), im Acceptable Noise Level Test (Nabelek *et al.*, 1991, 2006) und im Perceptual Performance Test (Saunders *et al.*, 2004; Saunders und Cienkowski, 2002). In allen Tests wird Sprache zusammen mit einem Hintergrundgeräusch präsentiert und ein Signal-Rausch-Abstand bestimmt. Eine andere perzeptive Untersuchung ist der Paarvergleich (Bortz und Döring, 2005), mit dem durch paarweisen Vergleich zwei Stimuli hinsichtlich eines Qualitätskriteriums beurteilt werden. Die Tests sollen überprüfen, ob Versuchspersonen bei Verwendung von Algorithmen zur Störgeräuschreduktion mehr Hintergrundgeräusch tolerieren als in Situationen ohne Störgeräuschreduktion oder ob eine Verbesserung der Klangqualität durch die Störgeräuschreduktion erreicht werden kann.

Technische Maße zur Nutzendokumentation von Störgeräuschreduktionen sind z.B. die Modulationstransferfunktion (Drullmann *et al.*, 1994) und der Δ SNR nach Hagerman und Olofsson (2004). Diese Methoden berechnen Werte, die einen Hinweis auf die Veränderung des Gesamtsignals durch einen Störgeräuschreduktionsalgorithmus geben.

Kap. 2 beschreibt die Methoden der einkanaligen Störgeräuschreduktion (s. Kap. 2.3). Eine Auswahl perzeptiver und technischer Maße, die zur Dokumentation des Nutzens der Störgeräuschreduktionsalgorithmen verwendet werden können, wird in Kap. 2.4 und Kap. 2.5 vorgestellt.

2.3 Einkanalige Störgeräuschreduktionen

Ziel der Verfahren zur Störgeräuschreduktion (engl. *Noise Reduction oder Speech Enhancement*) ist, die Störung des Nutzsignals s durch ein Störsignal n mit Hilfe von adaptiven Signalverarbeitungsalgorithmen möglichst weitgehend zu reduzieren und dabei möglichst geringe Signalverzerrung zu erzeugen (Vary *et al.*, 1998). Als Nutzsignal wird in den meisten Fällen Sprache angenommen. Um die Zielsetzung zu erreichen, werden verschiedene Ansätze verwendet. Sie werden unterschieden nach der Anzahl der verwendeten Mikrofone in Ein-, Zwei- oder Mehrkanal-Ansätze. Das bedeutet, dass den Signalverarbeitungsalgorithmen zur Nutzsignalverbesserung bei einem einkanaligen Ansatz das vermischte Signal $x(k) = s(k) + n(k)$ zur Verfügung

steht, während beim zweikanaligen Ansatz die Eingangssignale $x_1(k) = s_1(k) + n_1(k)$ und $x_2(k) = s_2(k) + n_2(k)$ und bei einem mehrkanaligen Ansatz mit n Kanälen $x_1(k) = s_1(k) + n_1(k)$, $x_2(k) = s_2(k) + n_2(k)$, ..., $x_n(k) = s_n(k) + n_n(k)$ anliegen. Die unterschiedlichen Eingangssignale führen zu unterschiedlichen Strategien der Störgeräuschreduktion. Diese Strategien basieren auf der Unterscheidung zwischen dem Nutz- und dem Störsignal. Im einkanaligen Ansatz können dabei nur durch Analyse statistischer und spektraler Unterschiede die Signalkomponenten von x differenziert werden. Die verwendeten einkanaligen Ansätze analysieren zuerst das Störsignal (Martin, 1993, 1994, 2001, 2003; Cohen und Berdugo, 2002; Tchorz und Kollmeier, 2003; Appiah *et al.*, 2005). Die gewonnenen Informationen werden im Folgenden zur Reduktion verwendet. Klassische Ansätze sind die Optimalfilterung, Spektrale Subtraktion und abgeleitete Varianten (Ephraim und Malah, 1984, 1985; Vary *et al.*, 1998).

In dieser Arbeit sollen einkanalige Störgeräuschreduktionen mit unterschiedlichen perceptiven und technischen Methoden untersucht werden. Die Grundlagen und Strategien der Störgeräuschreduktionen werden im Folgenden genauer beschrieben.

2.3.1 Optimalfilter (Wienerfilter)

Abb. 2.1 zeigt, dass der einkanaligen Störgeräuschreduktion nur das gestörte Signal $x(k)$ zur Verfügung steht. Es entsteht durch additive Überlagerung $x(k) = s(k) + n(k)$ des Nutzsignals $s(k)$ mit dem Störsignal $n(k)$.

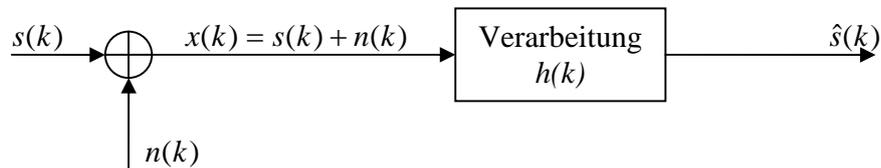


Abb. 2.1: Filterung eines gestörten Signals zur Störgeräuschreduktion

Vorausgesetzt wird, dass es sich bei der Filterung zur Störgeräuschreduktion um mittelwertfreie und stationäre Prozesse handelt. Gesucht wird eine Impulsantwort $h(k)$ bzw. eine Übertragungsfunktion $H(z)$, mit der das gestörte Signal $x(k)$ gefiltert wird. Ergebnis ist das geschätzte Signal $\hat{s}(k)$, das als störsignalbefreites Nutzsignal beschrieben und durch die Faltungssumme in Formel 2.1 berechnet werden kann. $h(k)$ bzw. dessen Fouriertransformierte $H(z)$ sind so zu wählen, dass das geschätzte Nutzsignal $\hat{s}(k)$ möglichst genau dem tatsächlichen Nutzsignal $s(k)$ entspricht. Geschätzte Signale werden im Folgenden mit einem Dach (z. B. \hat{s}) gekennzeichnet.

$$\hat{s}(k) = \sum_{\kappa=-\infty}^{\infty} h(\kappa)x(k - \kappa) = h(k) * x(k) \quad (2.1)$$

Der Fehler $e(k)$ zwischen dem ungestörten Nutzsignal $s(k)$ und dem Filterausgang $\hat{s}(k)$ wird durch Subtraktion berechnet:

$$e(k) = \hat{s}(k) - s(k) \quad (2.2)$$

Als Kriterium für eine gute Annäherung des geschätzten an das tatsächliche Nutzsignal wird der minimale mittlere quadratische Fehler (engl. *minimum mean-square error*, *MMSE*) verwendet:

$$\overline{e^2} = E \{e^2(k)\} = E \left\{ (\hat{s}(k) - s(k))^2 \right\} \stackrel{!}{=} \min \quad (2.3)$$

Durch Einsetzen von 2.1 in 2.3 entsteht die Bedingung:

$$\overline{e^2} = E \left\{ \left(\sum_{\kappa=-\infty}^{\infty} h(\kappa)x(k-\kappa) - s(k) \right)^2 \right\} \quad (2.4)$$

Zur Minimierung des quadratischen Fehlers $\overline{e^2}$ müssen die partiellen Ableitungen nach $h(i)$ gleich Null gesetzt werden (Vary *et al.*, 1998):

$$\frac{\partial \overline{e^2}}{\partial h(i)} = 0 \quad (2.5)$$

Das Ergebnis ist ein lineares Gleichungssystem der Form

$$\sum_{\kappa=-\infty}^{\infty} h(\kappa)\varphi_{xx}(i-\kappa) = \varphi_{xs}(i) \quad (2.6)$$

mit der Autokorrelationsfunktion φ_{xx} des gestörten Signals und der Kreuzkorrelationsfunktion φ_{xs} zwischen dem gestörten Signal und dem ungestörten Nutzsignal. Diese allgemeine Beziehung ist Grundlage für die Frequenzbereichslösung, die durch Fouriertransformation erreicht wird:

$$\sum_{\kappa=-\infty}^{\infty} h(\kappa)\varphi_{xx}(i-\kappa) = \varphi_{xs}(i) \quad \circ \longrightarrow \bullet \quad H(e^{j\omega})\Phi_{xx}(e^{j\omega}) = \Phi_{xs}(e^{j\omega}), \quad (2.7)$$

worin $\Phi_{xx}(e^{j\omega})$ das Autoleistungsdichtespektrum des gestörten Signals und $\Phi_{xs}(e^{j\omega})$ das Kreuzleistungsdichtespektrum des gestörten Signals und des ungestörten Nutzsignals sind. Ein optimaler Filter hat nach 2.7 den Frequenzgang:

$$H(e^{j\omega}) = \frac{\Phi_{sx}(e^{j\omega})}{\Phi_{xx}(e^{j\omega})} = \frac{\Phi_{ss}(e^{j\omega}) + \Phi_{ns}(e^{j\omega})}{\Phi_{ss}(e^{j\omega}) + \Phi_{sn}(e^{j\omega}) + \Phi_{ns}(e^{j\omega}) + \Phi_{nn}(e^{j\omega})} \quad (2.8)$$

Wird vorausgesetzt, dass das Nutzsignal $s(k)$ und das Störsignal $n(k)$ unkorreliert sind, werden die Kreuzkorrelationen zwischen Nutz- und Störsignal Null:

$$H(e^{j\omega}) = \frac{\Phi_{ss}(e^{j\omega})}{\Phi_{ss}(e^{j\omega}) + \Phi_{nn}(e^{j\omega})} \quad (2.9)$$

Diese Filterfunktion wird auch „Wienerfilter“ oder Optimalfilter genannt. Sie ist eine theoretische Optimallösung für die Rauschunterdrückung und setzt voraus, dass Nutz- und Störsignal bekannt sind.

2.3.2 Spektrale Subtraktion und ihr Zusammenhang mit dem Optimalfilter

Der Ansatz der Spektralen Subtraktion geht wie der Optimalfilteransatz davon aus, dass Nutz- und Störsignal unkorreliert sind. Daraus folgt, dass sich die Leistungsdichtespektren beider Signale zum Leistungsdichtespektrum des gestörten Signals addieren (Vary *et al.*, 1998):

$$\Phi_{xx}(e^{j\omega}) = \Phi_{ss}(e^{j\omega}) + \Phi_{nn}(e^{j\omega}), \quad (2.10)$$

woraus durch Umstellung nach $\Phi_{ss}(e^{j\omega})$ folgt

$$\Phi_{ss}(e^{j\omega}) = \Phi_{xx}(e^{j\omega}) - \Phi_{nn}(e^{j\omega}). \quad (2.11)$$

Durch Ausklammern des Faktor $\Phi_{xx}(e^{j\omega})$ lässt sich die Spektrale Subtraktion in 2.11 als Produkt beschreiben.

$$\Phi_{ss}(e^{j\omega}) = \Phi_{xx}(e^{j\omega}) \left[1 - \frac{\Phi_{nn}(e^{j\omega})}{\Phi_{xx}(e^{j\omega})} \right] \doteq \Phi_{xx}(e^{j\omega}) \left| \tilde{H}(e^{j\omega}) \right|^2 \quad (2.12)$$

Die Differenz in den eckigen Klammern in 2.12 beschreibt den Betragsfrequenzgang $\left| \tilde{H}(e^{j\omega}) \right|^2$ eines Filters, das auf das gestörte Signal $x(k)$ wirkt.

$$\tilde{H}(e^{j\omega}) = \sqrt{1 - \frac{\Phi_{nn}(e^{j\omega})}{\Phi_{xx}(e^{j\omega})}} \quad (2.13)$$

Basierend auf psychoakustischen Experimenten kann die Subtraktionsregel nach 2.12 verallgemeinert werden (Vary *et al.*, 1998). Als „bereinigtes Spektrum“ wird dann definiert

$$\begin{aligned} \hat{\Phi}_{ss}(e^{j\omega}) &\doteq \Phi_{xx}(e^{j\omega}) \left[1 - \left(\frac{\Phi_{nn}(e^{j\omega})}{\Phi_{xx}(e^{j\omega})} \right)^\epsilon \right]^\delta \\ &= \Phi_{xx}(e^{j\omega}) \left| \tilde{H}(e^{j\omega}) \right|^2 \end{aligned} \quad (2.14)$$

mit dem Filter $\tilde{H}(e^{j\omega})$, das „entstörend“ wirkt. Durch eine geeignete Wahl von δ und ϵ können Sonderfälle definiert werden:

- Ist $\delta = 2$ und $\epsilon = \frac{1}{2}$ wird 2.14 zu

$$\tilde{H}(e^{j\omega}) = 1 - \sqrt{\frac{\Phi_{nn}(e^{j\omega})}{\Phi_{xx}(e^{j\omega})}}. \quad (2.15)$$

In der Spektralen Subtraktion werden nun die Betragsspektren subtrahiert.

$$\sqrt{\hat{\Phi}_{ss}(e^{j\omega})} = \sqrt{\Phi_{xx}(e^{j\omega})} - \sqrt{\Phi_{nn}(e^{j\omega})} \quad (2.16)$$

- Ist $\delta = \epsilon = 1$ kann die allgemeine Form 2.14 zur spektralen Leistungsdichtesub-

traktion nach 2.13 umgeformt werden zu

$$\hat{\Phi}_{ss}(e^{j\omega}) = \Phi_{xx}(e^{j\omega}) - \Phi_{nn}(e^{j\omega}). \quad (2.17)$$

Um den Zusammenhang zwischen der Spektralen Subtraktion und dem Optimalfilter darzustellen, können die Frequenzgänge $H(e^{j\omega})$ verglichen werden. Für den Optimalfilter gilt nach 2.9 und durch Substitution mit 2.10

$$H_{opt}(e^{j\omega}) = \frac{\Phi_{ss}(e^{j\omega})}{\Phi_{ss}(e^{j\omega}) + \Phi_{nn}(e^{j\omega})} = \frac{\Phi_{ss}(e^{j\omega})}{\Phi_{xx}(e^{j\omega})}. \quad (2.18)$$

Die Spektrale Subtraktion ist beschrieben in 2.13 und kann umgeformt werden zu

$$\tilde{H}_{SpekSub}(e^{j\omega}) = \sqrt{1 - \frac{\Phi_{nn}(e^{j\omega})}{\Phi_{xx}(e^{j\omega})}} = \sqrt{\frac{\Phi_{ss}(e^{j\omega})}{\Phi_{xx}(e^{j\omega})}} \quad (2.19)$$

Damit ist der Zusammenhang zwischen Spektraler Subtraktion und Optimalfilter:

$$\tilde{H}_{SpekSub}(e^{j\omega}) = \sqrt{H_{opt}(e^{j\omega})} \quad (2.20)$$

2.3.3 Realisierung des Optimalfilters und der Spektralen Subtraktion

Zur Realisierung des Optimalfilters und der Spektralen Subtraktion werden die Signale in den Frequenzbereich transformiert und die Vorteile der schnellen Fouriertransformation (engl. *Fast Fourier Transformation, FFT*) genutzt. Die oben beschriebenen Algorithmen basieren auf der Annahme, dass die Signale stationär sind. Diese Annahme kann für ein Sprachsignal nicht aufrechterhalten werden. Sprache gilt jedoch als kurzzeitstationär. Mit Hilfe der Weighted Overlap-Add-Methode werden die Signale deshalb in Blöcken verarbeitet, für die die Annahme der Stationarität gilt (Vary *et al.*, 1998). Die Filterung 2.1 wird im Zeitbereich durch eine Faltung bzw. nach der Transformation der Signale im Frequenzbereich als Multiplikation ausgedrückt:

$$\hat{s}(k) = \sum_{\kappa=-\infty}^{\infty} h(\kappa)x(k - \kappa) = h(k) * x(k) \quad \bullet \rightarrow H(n)X(n), \quad (2.21)$$

wobei n der Frequenzindex ist.

Wie in Kap. 2.3.1 und Kap. 2.3.2 beschrieben, müssen für die Bildung des Filters zur Störgeräuschreduktion die Leistungsdichtespektren des reinen Störsignals Φ_{nn} und des gemischten Signals Φ_{xx} bekannt sein. Die Schätzung von Φ_{nn} wird in Kap. 2.3.5 beschrieben.

Die Annahme der Kurzzeitstationarität und die damit verbundene Verarbeitung der Signale in Blöcken benötigt die Anpassung der Leistungsdichtespektren über die gesamte Signallänge. Die Anpassung wird über den Glättungsparameter α gesteuert. Er sollte umso größer gewählt werden, je stationärer der Prozess ist (Kroschel, 2004). Für die Spektrale Subtraktion ergibt sich folgende Schätzung der Leistungsdichtespektren $\hat{\Phi}_{ss}$

und $\hat{\Phi}_{nn}$:

$$\hat{\Phi}_{nn}(n, \ell) = \alpha_{nn} \cdot \hat{\Phi}_{nn}(n, \ell - 1) + (1 - \alpha_{nn}) \cdot |N(n, \ell)|^2 \quad (2.22)$$

$$\hat{\Phi}_{xx}(n, \ell) = \alpha_{xx} \cdot \hat{\Phi}_{xx}(n, \ell - 1) + (1 - \alpha_{xx}) \cdot |X(n, \ell)|^2 \quad (2.23)$$

Auf ähnliche Weise kann eine Adaptation auch in der Optimalfilterung berücksichtigt werden (Kroschel, 2004).

Der in 2.22 und 2.23 verwendete Glättungsparameter α kann als Zeitkonstante τ angegeben werden. Der Zusammenhang zwischen α und τ ist durch

$$\alpha = \exp\left(\frac{-d}{\tau \cdot fs}\right) \quad (2.24)$$

gegeben. Darin ist d der Vorschub der Blöcke. Er ist definiert als Differenz zwischen der Blocklänge und der Länge der Überlappung in der Blockverarbeitung.

Zur Verbesserung der Spektralen Subtraktion können zwei Konstanten verwendet werden (Kroschel, 2004):

1. Der **Überschätzfaktor** a steuert den Einfluss des geschätzten Leistungsdichtespektrums des Rauschsignals $\hat{\Phi}_{nn}$ im Filter. Ist diese Schätzung gut, kann der Wert klein gewählt werden. Typisch sind Werte zwischen $a = 1$ und $a = 4$.
2. Der **Spectral Floor** b ermöglicht eine Verbesserung des subjektiven Höreindrucks. Im Fall, dass das geschätzte Leistungsdichtespektrum des Rauschens größer ist als das des gestörten Signals, können durch die Subtraktion negative Werte entstehen. Dies widerspricht der Definition des Leistungsdichtespektrums. Die negativen Werte werden durch 0 oder einen Wert b ersetzt. Ist $b > 0$ entsteht ein Rauschteppich, der den subjektiven Höreindruck verbessert.

Werden die genannten Parameter a und b verwendet, ergibt sich für die Spektrale Subtraktion:

$$\hat{H}(n) = P \left[\sqrt{1 - a \cdot \frac{\hat{\Phi}_{nn}(n, \ell)}{\hat{\Phi}_{xx}(n, \ell)}} \right] \quad (2.25)$$

mit

$$P(x) = \begin{cases} x, & \text{für } x \geq b \\ b, & \text{sonst} \end{cases} \quad (2.26)$$

2.3.4 Störgeräuschreduktion nach Ephraim und Malah

Zur Darstellung des Verhältnisses zwischen Signalleistung und Rauschleistung wird das Signal-Rausch-Verhältnis (engl. *Signal-to-Noise Ratio*, *SNR*) angegeben. Im Zeitbereich ist dieser definiert als

$$SNR = \frac{\sigma_s^2}{\sigma_n^2} = \frac{\varphi_{ss}(0)}{\varphi_{nn}(0)} \quad (2.27)$$

mit den Autokorrelationsfunktionen des Nutz- und Störsignals zum Zeitpunkt $k = 0$, $\varphi_{ss}(0)$ und $\varphi_{nn}(0)$. Diese Definition gilt für mittelwertfreie Signale und wird im Spek-

tralbereich zu

$$SNR = \frac{\Phi_{ss}(e^{j\omega})}{\Phi_{nn}(e^{j\omega})} \quad (2.28)$$

basierend auf den Leistungsdichtespektren des Nutzsignals $\Phi_{ss}(e^{j\omega})$ und des Störsignals $\Phi_{nn}(e^{j\omega})$. Der Signal-Rausch-Abstand beschreibt den logarithmierten SNR in der Einheit dB

$$10 \cdot \log_{10}(SNR) \text{ in [dB]}. \quad (2.29)$$

Ephraim und Malah (1984, 1985) entwickelten einen Algorithmus, der störende Artefakte, die bei anderen einkanalen Störgeräuschreduktionen zurückbleiben, unterdrückt. Cappé (1994) hat diese verwendete Methode untersucht und zusammengefasst. Definiert werden in der Ephraim-Malah-Suppression-Rule (EMSR) nach Cappé (1994) ein *a posteriori SNR* und ein *a priori SNR*

$$S\hat{N}R_{post}(n, \ell) = \frac{\hat{\Phi}_{xx}(n, \ell)}{\hat{\Phi}_{nn}(n, \ell)} - 1 = \frac{\hat{\Phi}_{xx}(n, \ell) - \hat{\Phi}_{nn}(n, \ell)}{\hat{\Phi}_{nn}(n, \ell)} \quad (2.30)$$

$$S\hat{N}R_{prio}(n, \ell) = \frac{\hat{\Phi}_{ss}(n, \ell)}{\hat{\Phi}_{nn}(n, \ell)}. \quad (2.31)$$

Die Übertragungsfunktion der EMSR ist nach Ephraim und Malah (1984)

$$G(n) = \frac{\sqrt{\pi}}{2} \sqrt{\left(\frac{1}{1 + S\hat{N}R_{post}(n)} \right) \left(\frac{S\hat{N}R_{prio}(n)}{1 + S\hat{N}R_{prio}(n)} \right)} \cdot M \left[\left(1 + S\hat{N}R_{post}(n) \right) \left(\frac{S\hat{N}R_{prio}(n)}{1 + S\hat{N}R_{prio}(n)} \right) \right], \quad (2.32)$$

worin M die Funktion

$$M[\theta] = \exp\left(-\frac{\theta}{2}\right) \left[(1 + \theta) I_0\left(\frac{\theta}{2}\right) + \theta I_1\left(\frac{\theta}{2}\right) \right] \quad (2.33)$$

und I_0 und I_1 die modifizierten Bessel Funktionen nullter und erster Ordnung sind. Die Verstärkung $G(n)$ hängt von den beiden Parametern $S\hat{N}R_{post}$ und $S\hat{N}R_{prio}$ ab, die für jeden Block ℓ und jede Frequenz n in der Blockverarbeitung neu geschätzt werden. Zur Schätzung des $S\hat{N}R_{post}$ wird Gl. 2.30 verwendet. $S\hat{N}R_{prio}$ wird durch den sogenannten „decision directed approach“ ermittelt (Cappé, 1994):

$$\begin{aligned} S\hat{N}R_{prio}(n, \ell) &= (1 - \alpha)P\left\{S\hat{N}R_{post}(n, \ell)\right\} + \alpha \frac{\hat{\Phi}_{ss}(n, \ell - 1)}{\hat{\Phi}_{nn}(n, \ell - 1)} \\ &= (1 - \alpha)P\left\{S\hat{N}R_{post}(n, \ell)\right\} + \alpha \frac{|G(n, \ell - 1)X(n, \ell - 1)|^2}{\hat{\Phi}_{nn}(n, \ell - 1)} \end{aligned} \quad (2.34)$$

$P(\cdot)$ verhindert in 2.34 negative Schätzwerte.

$$P(x) = \begin{cases} x, & \text{für } x \geq 0 \\ 0, & \text{sonst} \end{cases} \quad (2.35)$$

Der Gewichtungskoeffizient α wurde experimentell bestimmt und sollte bei 0,98 liegen (Ephraim und Malah, 1984; Cappé, 1994).

Die ESMR berechnet einen *a posteriori* SNR (vgl. Gl. 2.30). Zusätzlich wird der *a priori* SNR mit Gl. 2.34 nichtlinear rekursiv geschätzt und die verwendeten Signalanteile gewichtet kombiniert. Aus beiden SNR-Werten wird dann mit Hilfe der Gl. 2.32 eine Abschwächung jedes Frequenzbins n ermittelt und auf das gestörte Signal angewandt.

Cappé (1994) zeigte, dass hauptsächlich der „decision directed approach“ zur Schätzung des *a priori* SNR die Verbesserung der Signalqualität verursacht, weniger die Verstärkungsfunktion $G(n)$. Ist SNR_{prio} klein, erfolgt eine starke Unterdrückung und umgekehrt. SNR_{post} wirkt als Korrekturparameter. Ist er klein (-20 dB), nähert sich die Verstärkungsfunktion $G(n)$ der Spektralen Subtraktion. Bei einem großen SNR_{post} (20 dB) ist sie vergleichbar mit dem Optimal-Filter. Reale Filter wählen spektrale Abschwächungen zwischen diesen Grenzfällen, hervorgerufen durch den SNR_{post} .

2.3.5 Methoden zur Störgeräuschschätzung

In allen bisher vorgestellten Algorithmen zur Störgeräuschreduktion wird neben dem Leistungsdichtespektrum des gestörten Signals $\hat{\Phi}_{xx}$ auch das Leistungsdichtespektrum des Störsignals $\hat{\Phi}_{nn}$ zur Berechnung des Filters benötigt. Wie bereits in Abb. 2.1 dargestellt, kann jedoch nur das gestörte Signal als bekannt vorausgesetzt werden. Unterschiedliche Ansätze werden verwendet, um das Störgeräusch zu schätzen. Unterschieden werden können z. B. die Methoden der Minimum Statistics, Voice Activity Detection mit Periodogrammschätzung, Amplitudenspektrogramme, und Minima Controlled Recursive Averaging zur Störgeräuschschätzung, die im Folgenden beschrieben werden.

Minimum Statistics: Die Methode der Störgeräuschschätzung, basierend auf Minimum Statistics, wurde von Rainer Martin entwickelt (Martin, 1993, 1994, 2001, 2003). Grundlage ist die Idee, dass die Maxima des gestörten Signals zur Sprache zugeordnet werden können und in den Minima das Störgeräusch geschätzt werden kann. In einem ersten Schritt wird die Leistungsdichte des gestörten Signals $\hat{\Phi}_{xx}$ geglättet.

$$\hat{\Phi}_{xx}(n, \ell) = \alpha \cdot \hat{\Phi}_{xx}(n, \ell - 1) + (1 - \alpha) \cdot |X(n, \ell)|^2 \quad (2.36)$$

Der Parameter α sollte in einem einfachen Ansatz zwischen 0,9 und 0,98 liegen (Martin, 1993, 1994). Er kann jedoch auch zeit- und frequenzabhängig ermittelt werden (Martin, 2001, 2003), um eine bessere Schätzung der Rauschleistung zu erlangen.

In einem zweiten Schritt wird dann die Rauschleistung, basierend auf dem geglätteten Leistungsdichtespektrum des gestörten Signals $\hat{\Phi}_{xx}$, geschätzt. Dazu wird das Minimum von $\hat{\Phi}_{xx}$ innerhalb eines Fensters der Länge D in Samples gesucht. Dieses Fenster ist unterteilt in W Sektoren der Länge M Samples, so dass gilt: $M \cdot W = D$. Die geringste Leistung des gestörten Signals wird durch einen Vergleich zwischen dem aktuellen Minimum und jedem Sample von $\hat{\Phi}_{xx}$ ermittelt. Nachdem M Samples verglichen wurden,

wird die minimale Leistung der letzten M Samples gespeichert und ein neuer Vergleichswert aus der aktuellen maximalen Leistung ermittelt. Um diese Minimumsuche an fluktuierende Störgeräusche anzupassen, sind weitere Arbeitsschritte erforderlich. Verändert sich die untersuchte minimale Leistung der letzten W Sektoren mit M Samples monoton, wird die geschätzte Rauschleistung schnell geändert. In diesem Fall entspricht die Rauschleistung dem Minimum der letzten M Samples. Verändern sich die gespeicherten Minima nicht monoton, wird das Minimum des letzten Fensters der Länge D als Rauschleistung angenommen. Ist die aktuelle geglättete Leistung des gestörten Signals $\hat{\Phi}_{xx}$ geringer als die geschätzte Leistung des Rauschens, wird diese Schätzung sofort angepasst. Sie entspricht dann dem Minimum der Leistung des gestörten Signals oder des aktuellen geschätzten Rauschsignals (Martin, 1993). Die so geschätzte Rauschleistung ist kleiner als die tatsächliche Rauschleistung, da in diesem Verfahren die Minima detektiert werden, und sie muss deshalb mit einem Korrekturfaktor gewichtet werden (engl. *Bias-Compensation*). Dieser ist zum einen abhängig von der Glättungskonstante α , die durch den SNR beeinflusst wird und zum anderen von der Fensterlänge D der Minimumschätzung. Ist der SNR klein, d. h. liegt nur Rauschen als Signal vor, nimmt α einen großen Wert an und die Leistungsdichte wird über einen langen Zeitraum geglättet. Ist der SNR groß, besteht das Signal aus Nutzsignal und Rauschen, α nimmt einen kleinen Wert an, und die Glättung des Signals geschieht über einen kurzen Zeitraum. So können nachfolgende Nutzsignalpausen erkannt werden und werden durch eine lange Glättung nicht verdeckt. Abhängig von der Glättungskonstanten α ist auch der Fehler, der bei der Minimumsuche entsteht und der durch die Bias-Compensation ausgeglichen werden muss. Diese Berechnung ist aufwendig und wird in Martin (2001) genauer beschrieben.

Amplitudenspektrogramme: Eine andere Möglichkeit die Leistung des Störgeräuschs zu schätzen, ist die Verwendung von Amplitudenmodulationsspektrogrammen (Tchorz und Kollmeier, 2003). Das gestörte Signal wird dazu in Muster transformiert, die neuro-physiologisch begründet sind. Die Amplitudenmodulationsspektrogramme enthalten Informationen über die Grundfrequenz und die Modulationsfrequenzen in einem Analysefenster. Sprache erzeugt charakteristische Amplitudenmodulationsspektrogramme, die sich von den Darstellungen der meisten Hintergrundgeräusche unterscheiden. Ein Neuronales Netzwerk wird verwendet, um zwischen den Spektrogrammen zu unterscheiden und den SNR der Signalanteile im gestörten Signal zu schätzen.

Für die Verarbeitung des gestörten Signals wird sein Langzeitpegel normalisiert und mit einer FFT in den Frequenzbereich transformiert. Die ermittelten Frequenzen werden in 15 Barkbändern zusammengefasst. Durch eine weitere FFT wird für jedes Frequenzband das Modulationsspektrum ermittelt. Auch die Modulationsspektren werden in 15 Kanälen zusammengefasst. Im Amplitudenmodulationsspektrogramm wird die Energie der Bandfrequenz in den Barkbändern über der Modulationsfrequenz aufgetragen. Es entsteht so eine Matrix von 15 x 15 Werten. Amplitudenmodulationsspektrogramme von Sprache zeigen neben der Grundfrequenz und den Harmonischen des Sprechers auch andere Modulationen der Sprache. Diese Muster können in einem Amplituden-

modulationsspektrogramm z. B. eines sprachsimulierenden Rauschens nicht dargestellt werden. Um Amplitudenmodulationsspektrogramme bei unterschiedlichen SNR-Werten zu erkennen, wurde ein Neuronales Netzwerk trainiert. Ergebnis einer Anwendung des Neuronalen Netzwerkes sind Schätzungen der SNR-Werte zwischen Sprache und Störgeräusch in 15 unterschiedlichen Kanälen. Diese Werte können zur Störgeräuschreduktion verwendet werden.

Voice Activity Detection und Periodogramme: Nach Appiah *et al.* (2005) ist die Aufgabe einer Voice Activity Detection (VAD), eine Aussage zu treffen über die Präsenz von Sprache in einem Signal und den Beginn und das Ende des Sprachsegmentes anzuzeigen. Dies ist bei unterschiedlichen Signalverarbeitungsstrategien von Vorteil. Z. B. kann bei einer Störgeräuschreduktion in den detektierten Sprachpausen die Leistung des Hintergrundgeräusches geschätzt werden. Ein VAD-Algorithmus extrahiert unterschiedliche Merkmale des Eingangssignals und vergleicht sie mit Schwellenwerten. Werden sie überschritten, so gibt der Algorithmus die Information zurück, dass Sprache im Signal vorliegt. Üblich ist in dieser Verarbeitung die Verwendung einer blockweisen Extraktion der Signalmerkmale und Klassifizierung. Verwendet werden z. B. die Messung der Spektralen Steigung, Autokorrelationen, das Verhältnis des log likelihood, Cepstral-, gewichtete Cepstral- und modifizierte Distanz-Messungen (Appiah *et al.*, 2005). Unabhängig von den verwendeten Methoden, besteht beim VAD immer der Kompromiss, dass Hintergrundgeräusche als Sprache oder Sprache als Hintergrundgeräusche detektiert werden.

In einer detektierten Sprachpause kann mit Hilfe eines Periodogramms im Frequenzbereich die Leistung des Hintergrundgeräusches geschätzt werden. Bei der Verwendung des Periodogramms wird berücksichtigt, dass der aktuelle Datenblock in der vorgenommenen Blockverarbeitung sich nicht periodisch in den folgenden Blöcken fortsetzt. Das Periodogramm einer Folge mit K Punkten ist definiert (Kammeyer und Kroschel, 2002) als

$$Per_K(e^{j\omega}) = \frac{1}{K} |X(e^{j\omega})|^2 = \frac{1}{K} X(e^{j\omega}) \cdot X^*(e^{j\omega}) \quad (2.37)$$

wobei X^* das konjugiert komplexe Spektrum ist. Das Periodogramm ist identisch mit der Fourier-Transformierten der nicht erwartungstreuen Schätzgröße für die Autokorrelationsfolge $\hat{\varphi}_{nn}$ (Kammeyer und Kroschel, 2002):

$$Per_K(e^{j\omega}) = \sum_{\kappa=-(K-1)}^{K-1} \hat{\varphi}_{nn}(\kappa) e^{-j\omega\kappa} \quad (2.38)$$

Die Wiener-Khintchine Beziehung verknüpft die Autokorrelationsfolge $\hat{\varphi}_{nn}$ und die spektrale Leistungsdichte $\hat{\Phi}_{nn}$ über die Fourier-Transformation:

$$\hat{\varphi}_{nn} \circ \bullet \hat{\Phi}_{nn} \quad (2.39)$$

Daher ergibt sich für das Leistungsdichtespektrum $\hat{\Phi}_{nn}$ folgende Schätzung:

$$\hat{\Phi}_{nn}(e^{j\omega}) = Per_K(e^{j\omega}) = \frac{1}{K} X(e^{j\omega}) \cdot X^*(e^{j\omega}) \quad (2.40)$$

Wird die Schätzung in einer blockweisen Verarbeitung vorgenommen, werden die Blockgrenzen durch eine rekursive Schätzung geglättet

$$\hat{\Phi}_{nn}(n, \ell) = \alpha \cdot \hat{\Phi}_{nn}(n, \ell - 1) + (1 - \alpha) \cdot |X(n, \ell)|^2, \quad (2.41)$$

wobei α der Glättungsparameter ist. Diese Schätzung der Rauschleistung wird nur in Sprachpausen vorgenommen. In Signalteilen, in denen Sprache und Rauschen vorliegen, wird die Schätzung angehalten und zuletzt geschätzte Werte verwendet.

Minima Controlled Recursive Averaging: Minima Controlled Recursive Averaging zur Störgeräuschschätzung wurde von Cohen und Berdugo (2002) entwickelt. Die Rauschleistung wird aus gewichteten zurückliegenden Leistungswerten und einem Glättungsparameter $\tilde{\alpha}_d$ geschätzt, der abhängig ist von der Wahrscheinlichkeit der Präsenz von Sprache im Frequenzband. Sie wird ermittelt aus dem Verhältnis der lokalen Leistung des gestörten Signals und seinem Minimum in einem Zeitfenster der Länge N Samples. Darin wird ähnlich wie in der Minimum Statistics zuerst das Minimum der Gesamtsignalleistung in einem bestimmten Fenster der Länge L Samples gesucht. Es wird als Grundlage verwendet, um zwischen Signalteilen aus Sprache und Störgeräusch und alleinigem Störgeräusch zu unterscheiden.

$$S_r(n, \ell) = \frac{S(n, \ell)}{S_{min}(n, \ell)} \quad (2.42)$$

beschreibt das Verhältnis der lokalen Energie S des gestörten Signals und seines ermittelten Minimums S_{min} . Überschreitet dieses Verhältnis eine Schwelle δ wird angenommen, dass Sprache zusammen mit einem Hintergrundgeräusch vorliegt ($I = 1$). Bleibt das Verhältnis unterhalb der Schwelle, so enthalten die aktuellen Signalteile nur ein Hintergrundgeräusch ($I = 0$). Die Sprachwahrscheinlichkeit \hat{p} wird geschätzt über

$$\hat{p}(n, \ell) = \alpha_p \hat{p}(n, \ell - 1) + (1 - \alpha_p) I(n, \ell) \quad (2.43)$$

mit einem Glättungsparameter α_p , der zwischen 0 und 1 liegt. Schätzung der Sprachwahrscheinlichkeit \hat{p} wird in einem zeitveränderlichen Glättungsparameter $\tilde{\alpha}_d$ berücksichtigt

$$\tilde{\alpha}_d(n, \ell) = \alpha_d + (1 - \alpha_d) \hat{p}(n, \ell) \quad (2.44)$$

mit einem Glättungsparameter α_d zwischen 0 und 1. Das Leistungsdichtespektrum des Rauschens $\hat{\Phi}_{nn}$ wird dann geschätzt mit

$$\hat{\Phi}_{nn}(n, \ell + 1) = \tilde{\alpha}_d(n, \ell) \hat{\Phi}_{nn}(n, \ell) + (1 - \tilde{\alpha}_d(n, \ell)) |X(n, \ell)|^2. \quad (2.45)$$

Liegen Sprache und Rauschen vor, beträgt die Sprachwahrscheinlichkeit $\hat{p} = 1$. Nach 2.44 wird dann auch $\hat{\alpha}_d = 1$ und die Schätzung wird nach 2.45 angehalten. Dieser Vorgang ist vergleichbar mit der VAD.

2.4 Perzeptive Maße zur Untersuchung von Sprache im Störgeräusch

Zur Untersuchung des Nutzens von Störgeräuschreduktionsverfahren bieten sich verschiedene Untersuchungsformen von Sprache im Störgeräusch an, mit denen das Verhältnis von Sprache und Hintergrundgeräusch beschrieben wird. Die gemessenen Werte stellen dann durch einen Vergleich zwischen Situationen mit und ohne Störgeräuschreduktion den möglichen Nutzen dar. Eine Alternative zu diesen Messverfahren vergleicht die Signale mit und ohne Störgeräuschreduktion direkt, wie z. B. der Paarvergleich.

2.4.1 Acceptable Noise Level

Situationen, in denen Sprache im Störgeräusch vorkommt, können nicht nur durch die Sprachverständlichkeit, sondern auch durch die empfundene Klangqualität beschrieben werden. Hintergrundgeräusche in Sprachsituationen können z. B. einen großen Einfluss auf die Zufriedenheit mit Hörgeräten haben (Nabelek *et al.*, 1991). Eine Untersuchungsform von Sprache im Störgeräusch, die nicht nur das Sprachverstehen aufgreift, ist der Acceptable Noise Level Test (ANLT). Mit ihm wird die Frage beantwortet, wie viel Hintergrundgeräusch eine Versuchsperson überhaupt toleriert (Nabelek *et al.*, 1991).

Die Ermittlung des Acceptable Noise Level (ANL) erfolgt generell in drei Schritten. Zuerst wird der Pegel angenehmer Lautheit (engl. *Most Comfortable Level, MCL*) für Sprache bestimmt. Dazu wird der Pegel des Sprachsignals so lange geregelt, bis die Versuchsperson angibt, dass sie die Sprache angenehm empfindet und ihr über längere Zeit ohne Mühe folgen kann. In einer zweiten Messung wird ein tolerierbarer Pegel eines zusätzlichen Hintergrundgeräusches (engl. *Background Noise Level, BNL*) ermittelt. Dazu werden Sprache beim MCL und ein im Pegel veränderliches Hintergrundgeräusch präsentiert. Die Versuchsperson hat den maximalen Pegel des Geräusches zu bestimmen, bei dem die Sprache gut zu verstehen ist und sie dem Gesamtsignal über längere Zeit zuhören kann, ohne zu ermüden. Der so ermittelte Pegel des Rauschens ist der BNL. Der ANL wird dann errechnet aus der Subtraktion des BNL vom MCL. Abb. 2.2 zeigt den schematischen Ablauf der Ermittlung des ANL.

Mit dem ANLT wurden unterschiedliche Schwerpunkte und Fragestellungen untersucht. Nicht nur eine Beurteilung der Hörgerätenutzung und -zufriedenheit oder die Untersuchung unterschiedlicher Hörgerätemerkmale wurden durchgeführt, sondern auch das Testverfahren ANLT selbst wurde untersucht.

ANL und die Vorhersage der Hörgerätenutzung In einer Studie von Nabelek *et al.* (1991) wurde das Messverfahren ANLT zum ersten Mal verwendet und mit ihm der Zusammenhang zwischen der Hörgerätenutzung und der Toleranz von Hintergrundgeräuschen untersucht. Vollzeithörgeräteträger tolerierten signifikant höhere Pegel von Rauschen mit Sprachspektrum und Musik als Teilzeithörgeräteträger und Schwerhörige, die ihre Hörgeräte nicht nutzen. Außerdem bezeichneten sich Vollzeithörgeräteträger als

weniger beeinträchtigt mit ihren Hörgeräten in alltäglichen Situationen. Diese Aussage traf auf Teilzeithörgeräteträger nicht zu.

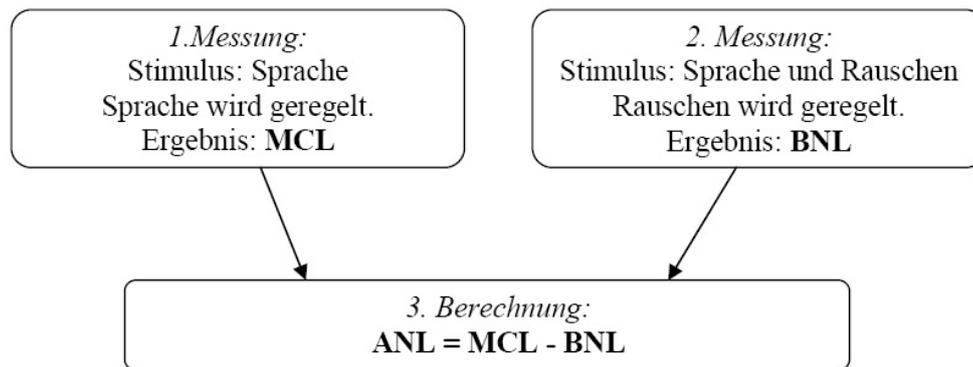


Abb. 2.2: Design des ANL

In einer weiteren Studie (Freyaldenhoven *et al.*, 2006a) wurde untersucht, ob mit Hilfe des ANL bei unterschiedlichen Sprachpegeln zwischen Vollzeit-, Teilzeit-Hörgeräteträgern und Nichtnutzern unterschieden werden und ihre Hörgerätenutzung vorausgesagt werden kann. Um die Ergebnisse auszuwerten, wurde ein globaler ANL (Mittelwert aller ermittelten ANL bei allen Sprachpegeln) und der ANL-Zuwachs (Steigung der ANL-Veränderung bei unterschiedlichen Startpegeln, ermittelt durch lineare Regression) errechnet. Es zeigte sich, dass der konventionell bestimmte ANL mit MCL als Sprachpegel, der globale ANL und der ANL-Zuwachs die Hörgeräte-Nutzung voraussagen können. Diese Ergebnisse konnten durch eine weitere Studie (Nabelek *et al.*, 2006) bestätigt werden. In ihr wurde nachgewiesen, dass mit Hilfe des unversorgten ANL eine Zuordnung der Versuchspersonen in zufriedene (Vollzeitnutzer) und unzufriedene (Teilzeit- und Nichtnutzer) Hörgeräteträger mit einer Sicherheit von 85% möglich ist. Daraus wurde verallgemeinert, dass Versuchspersonen mit einem ANL, der kleiner oder gleich 7 dB ist, zufriedene Hörgerätenutzer sind. Wählen sie einen ANL zwischen 7 und 13 dB ist keine Aussage über ihre Zufriedenheit mit den Hörgeräten möglich. Bei einem ANL über 13 dB sind die Versuchspersonen unzufriedene Hörgeräteträger.

ANL und der Nutzen verschiedener Hörgerätemerkmale Weitere Studien untersuchten den Nutzen verschiedener Hörgerätemerkmale mit dem ANLT. Freyaldenhoven *et al.* (2005) verglichen die Ergebnisse des ANLT mit Schwellen 50%iger Wortverständlichkeit im Hintergrundgeräusch (engl. *Speech Reception Thresholds, SRT*) und Front to Back Ratios (FBR), um den Richtungs-Nutzen von omnidirektionalen Hörgerätemikrofonen zu ermitteln. Dieser Richtungs-Nutzen wurde definiert als Differenz zwischen den ermittelten Werten bei omnidirektionaler und richtungs-Einstellung des Hörgerätemikrofons. Dabei wurde kein signifikanter Unterschied zwischen den Richtungs-Nutzen des ANL, SRT und FBR ermittelt.

Freyaldenhoven *et al.* (2006e) prüften den Effekt von Low-Frequency Gain Compensation und Venting bei Hörgeräten. Richtungs-Hörgeräte erzeugen geringere Verstärkung in den tiefen Frequenzen als Hörgeräte mit omnidirektionalen Mikrofonen. Tiefe Si-

gnale sind in der Phase ähnlicher als hohe, wenn sie den Mikrofoneingang erreichen und werden durch direktionale Mikrofone gedämpft. Der Verlust der Verstärkung für tiefe Frequenzen kann durch eine Low-Frequency Gain Compensation ausgeglichen werden. Als Untersuchungsverfahren wählten Freyaldenhoven *et al.* (2006e) FBR, SRT und ANLT. Die Ergebnisse zeigten, dass die Größe des Ventings verändert werden konnte, ohne die Sprachverständlichkeit oder die Akzeptanz von Hintergrundgeräuschen zu beeinflussen. Die Low Frequency Gain Compensation kann zu einer Verbesserung des direktionalen Nutzens bei Versuchspersonen mit einem tieffrequenten Hörverlust von mindestens 55 dB HL führen.

Außerdem wurde von Freyaldenhoven *et al.* (2006d) die Akzeptanz von Hintergrundgeräuschen mit monauraler und binauraler Verstärkung durch Hörgeräte untersucht (Freyaldenhoven *et al.*, 2006d). Das Sprachverstehen nimmt bei binauraler Verstärkung zu. Hingegen gilt für die meisten Versuchspersonen, dass der ANL nicht von der Art der Verstärkung abhängig ist. Bei einigen Versuchspersonen resultierte eine monaurale Verstärkung in einem größeren ANL. Außerdem wurde festgestellt, dass interaurale Unterschiede des ANL auftreten können.

Um den Nutzen einer digitalen Störgeräuschreduktion eines Hörgerätes zu prüfen, verwendeten Mueller *et al.* (2006) den ANLT und einen Sprachverständlichkeitstest im Störgeräusch. Beide Messungen wurden mit und ohne Störgeräuschreduktion durchgeführt. Wurde die Störgeräuschreduktion eingeschaltet, nahmen die ANL-Werte signifikant zu. Gleiches konnte für die Ergebnisse des Sprachverständlichkeitstests nicht beobachtet werden.

Andere Studien mit dem ANLT untersuchten, wie sich der ANL bei einem konstanten Sprachpegel verhält und wie reliabel das Testverfahren ist. Auch ein Vergleich zwischen einem Sprachverständlichkeitstest und dem ANLT wurde durchgeführt. Franklin *et al.* (2006) untersuchten den Effekt des Sprachpegels bei der Bestimmung des ANL. Dazu wurden ANL-Messungen mit Normalhörenden bei festen Sprachpegeln verwendet. Festgestellt wurde, dass der ANL mit zunehmendem Pegel der Sprache zunimmt. Der ANL ist damit kein Einzelwert oder eine Konstante, sondern eine Vielzahl von Werten, die vom Sprachpegel abhängen. In einer weiterführenden Studie wurde der Zusammenhang zwischen Sprachpegel und ANL auch bei Schwerhörigen geprüft (Freyaldenhoven *et al.*, 2006b) und ermittelt, dass die Effekte des Startpegels auf den ANL unabhängig vom Hörverlust sind.

Reliabilität des ANL und Einfluss unterschiedlicher Hintergrundgeräusche

Außerdem wurde die Test-Retest-Reliabilität des ANL und seine Abhängigkeit von unterschiedlichen Hintergrundgeräuschen untersucht (Smiley *et al.*, 2006). Verwendet wurden Rauschen mit Sprachspektrum und Stimmengewirr. Die Ergebnisse zeigten, dass die ANL-Werte über einen kurzen Zeitraum reliabel sind, unabhängig von den verwendeten Hintergrundgeräuschen. Außerdem waren die gemittelten ANL-Werte abhängig vom Hintergrundgeräusch. Wie zuvor schon beschrieben, verwendete Nabelek *et al.* (1991) unterschiedliche Hintergrundgeräusche. Durch diese Studie wurde vor allem deutlich, dass Musik zu unterschiedlichen Ergebnissen führen kann. Dabei ist auch die Art

der Musik zu berücksichtigen. Diese Ergebnisse wurden von Gordon-Hickey und Moore (2007) aufgegriffen. Sie untersuchten den Einfluss von Musik und Musikgeschmack auf den ANL bei Normalhörenden und kamen zu dem Ergebnis, dass der ANL nicht mit dem Musikgeschmack zusammenhängt. ANL-Werte gemessen mit Stimmengewirr liegen niedriger als mit Musik gemessene Werte.

Vergleich des ANLT mit Sprachverständlichkeitstests Nabelek *et al.* (2004) verglichen die Ergebnisse des ANLT mit den Ergebnissen des Speech Perception in Background Noise Test (SPIN) Kalikov *et al.* (1977); Bilger *et al.* (1984), einem Sprachverständlichkeitstest. Untersucht wurden die Reliabilität der Tests und die Gewöhnung an die Testverfahren. Die Messungen wurden mit und ohne Hörgeräte durchgeführt. Die Ergebnisse zeigten, dass der ANLT ähnlich reliabel ist wie der SPIN. Ein Zusammenhang zwischen den Ergebnissen des ANLT und des SPIN konnte jedoch nicht gefunden werden. Die Ergebnisse des SPIN nehmen mit der Verstärkung zu. Dieses Verhalten zeigt der ANL nicht.

Da die Ergebnisse des ANLT sehr stark von den Versuchspersonen abhängig sind (Nabelek *et al.*, 1991; Crowley und Nabelek, 1996), wurde von Rogers *et al.* (2003) geprüft, ob dies durch das Geschlecht der Versuchspersonen erklärt werden kann. Festgestellt wurde, dass Männer zwar einen höheren MCL und BNL einstellten, aber die ANL-Werte nicht signifikant unterschiedlich zu den Ergebnissen von Frauen waren. Daraus wurde geschlossen, dass beide Geschlechter einen ähnlichen SNR zum Hören von Sprache im Störgeräusch bevorzugen.

Generell gibt Nabelek (2005) an, dass der ANL bei Versuchen mit 221 Normalhörenden zwischen -2 und 38 dB liegt mit einem Mittelwert bei 10 bis 11 dB. Versuche mit 315 Schwerhörigen zeigten eine ANL-Spanne von -2 bis 29 dB und einen Mittelwert bei 10 bis 11 dB.

2.4.2 Sprachverständlichkeitstests im Störgeräusch

Der Oldenburger Satztest (OLSA) ist ein mögliches Verfahren zur Untersuchung von Sprachverstehen im Störgeräusch (Wagner *et al.*, 1999a,b,c; Wagner, 2003). Seine Sätze besitzen die gleiche syntaktische Struktur: *Name Verb Zahl Adjektiv Objekt*. Je zehn Wörter stehen in diesen syntaktischen Gruppen zur Verfügung (s. Tab. 2.1), die unterschiedlich miteinander kombiniert werden können. So sind für den OLSA 40 Testlisten mit je 30 Sätzen entstanden. Außerdem gehört zum Testmaterial ein Rauschen, das durch zufällige Überlagerung des Sprachmaterials erzeugt wurde und dadurch das gleiche Spektrum wie die Sprache besitzt.

Den Versuchspersonen werden die OLSA-Sätze zusammen mit dem Rauschen vorgespielt und ihre Aufgabe ist es, die verstandenen Wörter zu wiederholen. In Abhängigkeit von den verstandenen Wörtern wird der Pegel der Sprache (oder auch des Rauschens) geregelt. Erzielt werden soll ein Pegel, bei dem die Versuchsperson 50% der Sprache versteht. Dieser Messwert wird Sprachverständlichkeitsschwelle im Hintergrundgeräusch (SRT) genannt.

Tab. 2.1: Design des OLSA. Die unterstrichenen oder kursiven Wörter zeigen Beispiele für verwendete Sätze (Wagener, 2003).

| Name | Verb | Zahl | Adjektiv | Objekt |
|--------------|----------------|-------------|--------------|----------------|
| Peter | <u>bekommt</u> | drei | große | <u>Blumen.</u> |
| Kerstin | sieht | <i>neun</i> | kleine | Tassen. |
| Tanja | kauft | sieben | alte | Autos. |
| Ulrich | gibt | acht | <i>nasse</i> | Bilder. |
| Britta | schenkt | <u>vier</u> | schwere | Dosen. |
| Wolfgang | verleiht | <u>fünf</u> | grüne | <i>Sessel.</i> |
| Stefan | hat | zwei | teure | Messer. |
| Thomas | gewann | achtzehn | schöne | Schuhe. |
| <i>Doris</i> | <u>nahm</u> | zwölf | <u>rote</u> | Steine. |
| <u>Nina</u> | <i>malt</i> | elf | weiße | Ringe |

Untersuchungen zeigten, dass mit dem OLSA adaptiv SRT-Werte gemessen werden können (Wagener *et al.*, 1999a,b; Wagener, 2003). Vor der Messung muss das Sprachmaterial geübt werden. Obwohl das Sprachmaterial kaum Redundanz besitzt, kann durch die geringe Anzahl der Wörter pro syntaktische Gruppe ein Lerneffekt auftreten (Wagener *et al.*, 1999a,b; Wagener, 2003). In den Untersuchungen können unterschiedliche Rauschen verwendet werden (Wagener, 2003).

2.4.3 Just Follow Conversation Test

Der Just Follow Conversation Test (JFCT) wurde in den späten 40'er Jahren zur Ermittlung des Sprachverstehens im Störgeräusch entwickelt (Hawkins und Stevens, 1950). Im JFCT wird den Versuchspersonen kontinuierliche Sprache zusammen mit einem Störgeräusch dargeboten. Im Gegensatz zu anderen Sprachverständlichkeitstests, verändern die Versuchspersonen selbständig den Pegel der Sprache (oder auch des Hintergrundgeräusches), bis sie die Sprache gerade eben verstehen. Dieser Vorgang unterliegt den subjektiven und individuellen Maßstäben der Versuchsperson, die als JFC-Kriterium bezeichnet werden. Zum selbständigen Einstellen des Pegels erhält die Versuchsperson ein Steuerungselement (z. B. Drehknopf oder Maus). Hat sie den Pegel eingestellt, so wird der SNR ermittelt und als JFC-Wert dokumentiert. Die Messungen werden vier bis zehnmal wiederholt und aus den letzten vier bis neun Pegeln der Median errechnet (Larsby und Arlinger, 1994; Borg *et al.*, 1998; Magnusson *et al.*, 2001).

Nach Neumann und Jensen (2001) können mit dem JFCT auf einfache und schnelle Weise subjektive Sprachverständlichkeitsschwellen ermittelt werden. Die Schwellen sind über kurze Zeit stabil, und es wird ein Sprachmaterial verwendet, das der natürlichen Sprachsituation sehr nahe kommt. Nachteilig ist, dass bei Wiederholung des Sprachmaterials ein Lerneffekt auftreten kann. Außerdem unterscheidet sich das JFC-Kriterium der Versuchspersonen, woraus interindividuelle Unterschiede entstehen. Die Ergebnisse können auch durch unterschiedliche oder undeutliche Instruktionen beeinflusst werden.

Eine Abwandlung des JFCT ist die HörTech Einpegelung der Individuellen Satzverständlichkeitsschwelle (HEISS). Abweichend vom JFC-Verfahren werden beim HEISS-

Test Sätze des Göttinger Satztestes (Kollmeier und Wesselkamp, 1997) verwendet, wobei ein Satz in ständiger Wiederholung mit nur kurzen Pausen dargeboten wird (Latzel *et al.*, 2003). Außerdem liegen der Versuchsperson die Test-Sätze in schriftlicher Form vor und können damit als bekannt vorausgesetzt werden.

Die Einstellung einer subjektiven Sprachverständlichkeitsschwelle wurde in unterschiedlichen Arbeiten zur Untersuchung von Störgeräuschreduktionen verwendet (Wittkop, 1994, 2001; Peissig, 1992; Peissig und Kollmeier, 1997). Der Pegel der Sätze, die mit einem Hintergrundgeräusch präsentiert wurden, wurde mit bestimmten Tasten in 1 dB-Schritten verändert, bis die Versuchspersonen 50% der Sätze subjektiv verstanden. Der Vorteil der Einregelung einer subjektiven Sprachverständlichkeitsschwelle gegenüber herkömmlichen Sprachtests ist der geringere zeitliche Aufwand. So ist es möglich mehrere Algorithmen und Einstellungen miteinander zu vergleichen.

2.4.4 Performance Perceptual Test

Saunders und Cienkowski (2002) beschrieben, dass die Hörerätzufriedenheit nicht nur aus verbesserter Sprachverständlichkeit besteht, sondern auch aus der Gesamtleistung des Hörgerätes, der Zufriedenheit des Höreräteträgers mit seinem Hörgerät und dem Gewinn, den er durch sein Hörgerät verspürt. Um die Zufriedenheit mit Hörgeräten zu ermitteln, forderten Saunders und Cienkowski ein Testverfahren, das Unterschiede zwischen subjektiven und objektiven Messgrößen untersucht und dabei das gleiche Testformat verwendet. So sollte ein Vergleich der Ergebnisse möglich werden.

Aus diesen Grundideen wurde der Performance Perceptual Test (PPT) entwickelt. Grundlage des PPT ist der Hearing-In-Noise-Test (HINT), ein englischer Sprachverständlichkeitstest. Mit dem im HINT verwendeten adaptiven Verfahren wird der SNR bestimmt, bei dem die Versuchsperson 50-%iges Satzverstehen erreicht. Dieser Messwert ist also mit dem Ergebnis des OLSA zu vergleichen.

Bei der Durchführung des PPT wird zuerst der Performance Speech Reception Threshold (Performance SRT) bestimmt. Dazu wird der HINT in seiner ursprünglichen Form durchgeführt und der SNR bestimmt, bei dem 50%-ige Satzverständlichkeit erreicht wird. In einer zweiten Messung wird die Aufgabe der Versuchsperson verändert. Ihr wird erneut der HINT vorgespielt, und sie soll den SNR wählen, bei dem sie meint, gerade eben die gesamte Sprache zu verstehen. Dieser Messwert ist der Perceptual SRT. Der JFCT beschreibt ein ähnliches Vorgehen, bei dem jedoch die Sätze kontinuierlich dargeboten werden und die Versuchsperson selbst auf den Pegel der Signale Einfluss nehmen kann. Ermittelt wird aber ebenfalls eine subjektive Verständlichkeit. Aus der Differenz zwischen dem Performance und Perceptual SRT wird die Performance Perceptual Discrepancy (PPDIS) bestimmt. Sie liefert einen Vergleich zwischen dem SNR, bei dem die Versuchsperson meint alle Sätze zu verstehen und dem SNR, der objektiv gemessen wurde. Eine positive PPDIS zeigt, dass eine Versuchsperson ihre Hörfähigkeit unterschätzt, eine negative PPDIS, dass die Versuchsperson ihre Hörfähigkeit überschätzt. Abb. 2.3 stellt den Ablauf des PPT dar.

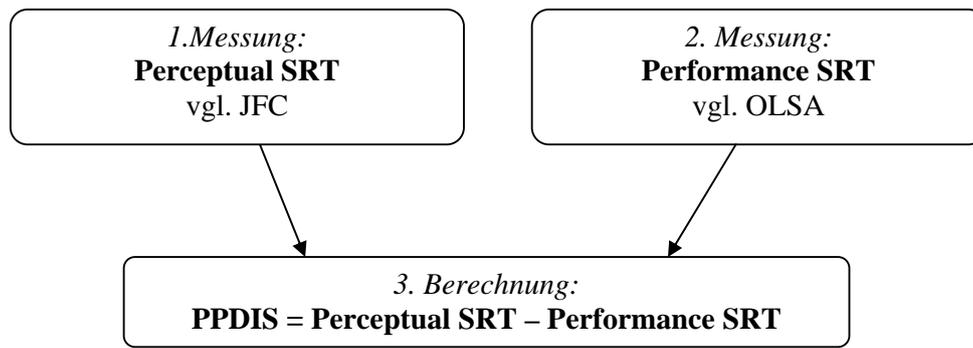


Abb. 2.3: Design des PPT.

In einer Studie von 2002 untersuchten Saunders und Cienkowski die subjektive und objektive Sprachverständlichkeit und ihren Zusammenhang zu beschriebenen Hörschwierigkeiten mit Hörgeräten, die mit einem Fragebogen ermittelt wurden. Die Ergebnisse zeigten, dass im Mittel 48% der Varianz in den subjektiven Bewertungen der Hörgeräte durch eine Kombination aus Performance SRT und PPDIS beschrieben werden konnten. Diese Untersuchung wurde durch Saunders *et al.* (2004) vertieft. Sie führten den PPT ohne Hörgeräte bei unterschiedlichen Sprachpegeln mit Schwerhörigen und Normalhörenden durch und ließen die Fragen des Hearing Handicap Inventory for Elderly/Adults (HHIE/A) (Ventry und Weinstein, 1982; Newman *et al.*, 1990) beantworten. Ein Unterschied zwischen den PPDIS-Werten konnte jedoch nicht ermittelt werden. Performance SRT, PPDIS und höheres Alter können 40% der Varianz des berichteten Handicaps (HHIE/A) erklären. Diese Untersuchungen wurden ebenso mit und ohne Hörgeräte durchgeführt (Saunders *et al.*, 2006). Versorgte und unversorgte PPDIS unterschieden sich nicht. PPDIS, Performance SRT und Alter sind signifikante Prädiktoren für Werte des HHIE/A, denn ein größeres empfundenes Handicap durch Schwerhörigkeit wird in Zusammenhang gebracht mit der Unterschätzung der Hörfähigkeit, geringem Alter und einer schlechteren Fähigkeit, mit Hörgeräten Sprache im Störgeräusch zu verstehen.

In einem Übersichtsartikel beschreiben Saunders *et al.* (2005) folgende Ergebnisse mit dem PPT: Die PPDIS korreliert signifikant mit der Dauer der Hörgerätenutzung, dem beschriebenen Nutzen der Hörgeräte im Hintergrundgeräusch, dem beschriebenen Nutzen für die Leichtigkeit der Kommunikation und dem beschriebenen positiven emotionalen Nutzen der Hörgeräte. Weder der Performance SRT noch der Nutzen der Hörgeräte, beschrieben durch die Differenz zwischen versorgtem und unversorgtem Performance SRT, korrelieren mit den o. g. Variablen. Die Messung des Hörgerätenutzens mit Sprachverständlichkeitstests ist nicht ausreichend, um die verschiedenen Nutzenaspekte der Hörgeräteversorgung zu beschreiben. Ein Vergleich der Ergebnisse mit denen des HHIE/A zeigt, dass Versuchspersonen mit einem geringen Handicap eine höhere PPDIS haben als Versuchspersonen mit einem hohen Handicap. Der Performance SRT beschreibt 8-10% der Varianz in der Hörgeräteeufriedenheit bei den berichteten Schwierigkeiten beim Hören mit Hörgerät und des Handicaps. PPDIS erklärt 23% der Varianz der Hörgerteeufriedenheit und zwischen 6-17% der Varianz der Hörschwierigkeiten und des Handicaps. Damit ist der beschriebene Hörgerätenutzen durch die Meinung über das eigene Hören und den aktuellen Nutzen, den Hörgeräte erzeugen, beeinflusst.

2.4.5 Paarvergleiche

Zur Beurteilung der Sprach- oder Signalqualität durch Versuchspersonen bieten sich verschiedene Verfahren an. Bortz und Döring (2005) nennen z. B. Rangordnungen, Dominanz-Paarvergleiche, Ähnlichkeits-Paarvergleiche und Rating-Skalen. Ein häufig gewähltes Verfahren ist der Dominanz-Paarvergleich. In ihm werden der Versuchsperson unterschiedliche Stimuli in Paaren präsentiert, von denen eines auszuwählen ist. Die Versuchsperson ist aufgefordert anzugeben, bei welchem von zwei Objekten das untersuchte Merkmal stärker ausgeprägt ist, bzw. welches Objekt bezüglich des Merkmals "dominiert". Das Kriterium, nach welchem die Auswahl zu treffen ist, wird durch den Versuchsleiter vorgegeben. Werden n Objekte ausgesucht, müssen für einen vollständigen Vergleich

$$\binom{n}{2} = \frac{n(n-1)}{2} \quad (2.46)$$

Urteile abgegeben werden. Das Ergebnis eines vollständigen Paarvergleiches besagt, wie oft ein Objekt den übrigen Objekten vorgezogen wurde. Um die Unterschiede zwischen den Paaren weiter zu untersuchen, kann zusätzlich eine Skalierung durchgeführt werden. Die Unterschiede zwischen zwei Objekten können dann z. B. mit den Stufen „sehr ähnlich“, „ähnlich“, „sehr unähnlich“ bewertet werden. Vorteil des Paarvergleiches ist, dass auch kleine Wahrnehmungsunterschiede zwischen den Objekten erfasst werden können. Das Messverfahren ist allerdings durch die Anzahl der Vergleiche sehr aufwändig und ermittelt lediglich ein Ergebnis auf Ordinalniveau.

Zur Auswertung des Dominanz-Paarvergleiches können die Konsistenz, d. h. die Widerspruchsfreiheit der Urteile einer Person, und die Konkordanz, das ist die Übereinstimmung der Urteile mehrerer Versuchspersonen, ermittelt werden (Bortz *et al.*, 2000). Sind Konsistenz und Konkordanz ausreichend, kann in einer ersten einfachen Auswertung aus diesen Ergebnissen eine indirekte Rangordnung ermittelt werden. Dazu wird die Häufigkeit ermittelt, mit der ein Objekt den anderen Objekten vorgezogen wurde und einem Rang zugeordnet. Ergebnis ist eine Rangfolge der verglichenen Stimuli (Kuk, 2002; Bortz und Döring, 2005). Eine einfache Form die Häufigkeiten zu ermitteln, ist die Darstellung der Ergebnisse in einer $n \times n$ Dominanzmatrix y_{ij} (s. Tab. 2.2). i ist dabei der Zeilenindex und j der Spaltenindex. Der Wert 1 in einer Zelle bedeutet, dass das Objekt in der gewählten Zeile i dem Objekt der gewählten Spalte j vorgezogen wurde. Der Wert 0 bedeutet dementsprechend, dass das Objekt in der gewählten Spalte j dem Objekt in der Zeile i vorgezogen wurde. Da nach 2.46 für einen vollständigen Paarvergleich lediglich der Vergleich $A - B$, aber nicht $B - A$ durchgeführt wird, werden in der Dominanzmatrix die entsprechenden Zellen mit der logischen Negation des Vergleiches $A - B$ ergänzt. Außerdem werden Objekte nicht mit sich selbst verglichen, und die Hauptdiagonale der Dominanzmatrix y_{ij} bleibt unbestimmt. Beurteilungen mehrerer Versuchspersonen werden ebenfalls in Matrizenform dargestellt und für ein Gesamturteil addiert. Daraus wird dann die relative Häufigkeitsmatrix p_{ij} berechnet, die zur Bildung der Rangreihenfolge genutzt werden kann.

Bei der bisher beschriebenen Auswertung des Paarvergleiches mit Rangreihenfolgen ist kein Rückschluss über die Abstände der Urteile möglich. Dazu muss eine Umformung

Tab. 2.2: Beispiel einer Dominanzmatrix für drei Stimuli A, B und C im Paarvergleich.

| | A | B | C |
|---|---|---|---|
| A | - | 1 | 1 |
| B | 0 | - | 0 |
| C | 0 | 1 | - |

der Ergebnisse mit Hilfe von psychologisch motivierten, mathematischen Modellen von ordinalskalierten Werten auf intervallskalierte Werte vorgenommen werden. Die Modelle gehen von der Voraussetzung aus, dass das zu beurteilende Merkmal eindimensional ist. Dies ist nach Sixtl (1967) dann gegeben, wenn der größte wahrnehmbare Unterschied zwischen drei Objekten gleich der Summe der beiden kleineren Unterschiede ist. Modelle, die angewendet werden können, sind z. B. Thurstones Law of Comparative Judgement oder das Bradley-Terry-Luce Modell (Edwards, 1957; Sixtl, 1967).

Zur Untersuchung der Qualität von Hörgeräten oder Hörgerätealgorithmen wurde der Paarvergleich von verschiedenen Autoren angewendet. Marzinzik (2000) konnte z. B. den Nutzen von verschiedenen Störgeräuschreduktionen in Hörgeräten mit Hilfe von Paarvergleichen nachweisen. Er untersuchte die Gesamtqualität, Natürlichkeit der Sprache, die Stärke der Störgeräuschreduktion und die Sprachverständlichkeit mit unterschiedlichen Hintergrundgeräuschen und verschiedenen SNR-Werten. Auch Ricketts und Hornsby (2005) konnten den Nutzen einer Störgeräuschreduktion in einem Hörgerät durch die Bewertung der Signalqualität im Paarvergleich nachweisen. Keinen signifikanten Nutzen durch eine Störgeräuschreduktion in einem Hörgerät konnten Boymans und Dreschler (2000) ermitteln.

2.5 Technische Maße zur Untersuchung von Sprache im Störgeräusch

Neben den in Kap. 2.4 beschriebenen perceptiven Maßen zur Untersuchung von Sprache im Störgeräusch können auch technische Maße verwendet werden. Sie vergleichen das Eingangs- mit dem Ausgangssignal einer Störgeräuschreduktion und dokumentieren auf diese Weise die SNR-Verbesserung.

2.5.1 SNR nach Hagerman und Olofsson

Hagerman und Olofsson (2004) entwickelten die nach ihnen benannte Untersuchung des SNR. Sie basiert auf der Idee, zwei unterschiedliche Messungen mit Sprache $s(t)$ und Störgeräusch $n(t)$ durchzuführen. Beide Messungen werden bei einem definierten SNR durchgeführt, wobei in einer das Rauschen invertiert wird

$$a_{in} = s(t) + n(t) \quad (2.47)$$

$$b_{in} = s(t) - n(t). \quad (2.48)$$

Diese Eingangssignale werden durch die Störgeräuschreduktion zu den Ausgangssignalen

$$a_{out} = s'(t) + n'(t) + e_1(t) \quad (2.49)$$

$$b_{out} = s'(t) - n'(t) + e_2(t), \quad (2.50)$$

wobei $s'(t)$ das Sprachsignal, $n'(t)$ das Rauschsignal und $e_1(t)$ bzw. $e_2(t)$ Fehlersignale am Ausgang der Störgeräuschreduktion sind. Die Fehlersignale entstehen durch die Verarbeitung der Störgeräuschreduktion. Aus den Ausgangssignalen werden die Signale $c(t)$ und $d(t)$ berechnet

$$c(t) = a_{out}(t) + b_{out} = 2s'(t) + e_1(t) + e_2(t) \quad (2.51)$$

$$d(t) = a_{out}(t) - b_{out} = 2n'(t) + e_1(t) - e_2(t). \quad (2.52)$$

Sind die Fehlersignale klein, so gelingt mit dieser Umrechnung eine Trennung des Sprach- und Rauschanteils des Ausgangssignals. Aus diesen Signalteilen können die Leistungsdichtespektren der Sprache $\Phi_{ss'}(f)$ und des Rauschens $\Phi_{nn'}(f)$ des Ausgangs geschätzt werden und mit den Leistungsdichtespektren $\Phi_{ss}(f)$ und $\Phi_{nn}(f)$ kombiniert werden, um den Betrag der Übertragungsfunktion der Störgeräuschreduktion für Sprache $|H_s(f)|^2$ und für Rauschen $|H_n(f)|^2$ zu berechnen

$$|H_s(f)|^2 = \frac{\Phi_{ss'}(f)}{\Phi_{ss}(f)} \quad (2.53)$$

$$|H_n(f)|^2 = \frac{\Phi_{nn'}(f)}{\Phi_{nn}(f)}. \quad (2.54)$$

Die Verbesserung des SNR durch die Störgeräuschreduktion wird mit dem Quotienten

$$\Delta SNR(f) = \frac{|H_s(f)|^2}{|H_n(f)|^2} \quad (2.55)$$

berechnet.

Um $\Delta SNR(f)$ als Zahlenwert zu beschreiben, können die SNR-Werte am Eingang und am Ausgang der Störgeräuschreduktion mit Hilfe des Speech Intelligibility Index (SII) durch Anwendung der Tabelle 1 der Norm ANSI S 3.5 - 1997 (1997) gewichtet werden. Die entstehenden skalaren Werte werden zu einem Quotienten ΔSNR zusammengefasst. Die gesamte Idee setzt voraus, dass Nutz- und Rauschsignal bekannt sind und dass die Störgeräuschreduktion quasi linear und zeitinvariant arbeitet. Fredelake (2006) wies den Nutzen einer Störgeräuschreduktion mit dem ΔSNR nach.

2.5.2 Modulationstransferfunktion

Durch eine Störgeräuschreduktion soll der SNR des gestörten Signals verbessert werden, wodurch sich die Modulationstiefe erhöht. Houtgast und Steeneken (1985) stellten die Modulationstransferfunktion (MTF) vor. Sie ist ein Maß, das die Veränderung der Modulationstiefe durch ein Übertragungssystem beschreibt. Angenommen, ein sinus-

förmiges Trägersignal $x_t(t)$ wird mit einer Frequenz f_{mod} und der Modulationstiefe m amplitudenmoduliert

$$x(t) = x_t(t)(1 + m \cdot \sin(2\pi f_{mod}t + \phi_0)). \quad (2.56)$$

Dieses Signal wird am Eingang eines Übertragungssystems angeboten und erzeugt ein Ausgangssignal der Form

$$y(t) = x_t(t)(1 + m' \cdot \sin(2\pi f_{mod}t + \phi_0)) \quad (2.57)$$

bei dem nur die Modulationstiefe m durch das System verändert wurde. Die MTF wird aus der Division der Modulationstiefe des Eingangs- und Ausgangssignals berechnet

$$MTF(f_{mod}) = \frac{m'}{m}. \quad (2.58)$$

Für beliebig modulierte komplexe Eingangssignale kann dieser Ansatz nach Houtgast und Steeneken (1985) nicht angewendet werden. Nach Drullmann *et al.* (1994) kann in diesem Fall die MTF mit Hilfe der Hilbertransformation ermittelt werden. Dazu wird die Division zwischen dem Kreuzleistungsdichtespektrum Φ_{xy} des Ein- und Ausgangssignals und dem Autoleistungsdichtespektrum Φ_{xx} des Eingangssignals berechnet

$$MTF_{komplex}(f_{mod}) = \alpha \cdot \frac{\Phi_{xy}(f_{mod})}{\Phi_{xx}(f_{mod})} \quad (2.59)$$

mit $\alpha = \frac{\Phi_{xy}(0)}{\Phi_{xx}(0)}$ als Normierungskonstante.

Die MTF_{PL} besteht aus dem Realteil der $MTF_{komplex}$ und berücksichtigt nur die Modulationen die phasensynchron zueinander sind.

$$MTF_{PL}(f_{mod}) = \text{Re} \{MTF_{komplex}(f_{mod})\} \quad (2.60)$$

Die Vereinfachung durch den Realteil ist unter der Annahme möglich, dass nur phasensynchrone Modulationen Einfluss auf die Sprachverständlichkeit haben. Der Index PL steht für „phase locked“ und beschreibt, dass nur Modulationen des Ausgangssignals, die in Phase sind mit den Modulationen des Eingangssignals, berücksichtigt werden. Durch das Übertragungssystem entstandene Modulationen bleiben so unberücksichtigt. Holube *et al.* (2005) untersuchten die Störgeräuschreduktionen verschiedener Hörgeräte mit der MTF. Die meisten Hörgeräte zeigten eine Verringerung des Faktors $\frac{1}{MTF_{PL}}$, was einer Beschreibung der Zunahme der Modulationstiefe durch die Übertragung im Hörgerät entspricht. Außerdem wies Fredelake (2006) den Nutzen einer Störgeräuschreduktion mit der MTF nach.

Kapitel 3

Untersuchung des Acceptable Noise Level Test

3.1 Zusammenfassung

In dieser Studie wurden verschiedene Fragestellungen zur Methode des Acceptable Noise Level Test (ANLT) in acht Experimenten untersucht. An jedem Experiment nahmen elf bis zwölf normalhörende Versuchspersonen teil. Die Ergebnisse zeigen, dass das von Nabelek et al. (1991, 2006) vorgestellte Verfahren ein zuverlässiges Maß zur subjektiven Beurteilung von Sprache im Störgeräusch ist. Es kann auf eine ähnliche Weise wie von Nabelek et al. (2006) vorgeschlagen, durchgeführt werden. Die verwendeten Schrittgrößen, Startpegel und Instruktionen können beibehalten werden. Außerdem zeigen die Acceptable Noise Level-Werte (ANL-Werte) eine hohe Test-Retest-Reliabilität. Werden mehrere Messungen in direkter Wiederholung durchgeführt, zeigen die ANL-Werte einen geringen Anstieg, hervorgerufen durch einen Anstieg des Most Comfortable Level (MCL). Aufgrund dieses Ergebnisses wurden ANL-Bestimmungen bei einem konstanten MCL durchgeführt, der einen Anstieg des ANL unterdrückte. Bei Wiederholungsmessungen sollte deshalb zuerst der MCL bestimmt und gemittelt, folgend der Background Noise Level gemessen werden. Generell kann ein Lerneffekt der Versuchspersonen durch Messwiederholungen berücksichtigt werden. Außerdem erreichten Versuchspersonen signifikant höhere Werte für ein Stimmengewirr von zwölf Sprechern als für andere Hintergrundgeräusche. Wird der ANLT in Zusammenhang mit einem Oldenburger Satztest durchgeführt, kann unberücksichtigt bleiben, ob der Speech Reception Threshold (SRT) vor oder nach dem ANL gemessen wird. Im letzten Experiment wurde der Einfluss der Darbietungspegel bei der Übung des Sprachmaterials auf den ANL analysiert. Ist eine Trainingsphase zur Übung des Sprachmaterials notwendig, spielt der Darbietungspegel für die Untersuchung des ANL keine Rolle. Für den Just Follow Conversation Test (JFCT) konnte aus den Experimenten gefolgert werden, dass die Messungen des JFC-Wertes nur in einem Termin und mit mindestens neun Messwiederholungen durchgeführt werden sollten. Weiterhin zeigten die Ergebnisse der Untersuchungen die Tendenz, dass der JFC-Wert durch eine SRT-Messung beeinflusst wird.

3.2 Einleitung

Der Acceptable Noise Level Test (ANLT) wurde von Nabelek *et al.* (1991) entwickelt und ermittelt das akzeptierbare Signal-Rausch-Verhältnis (engl. *Signal-to-Noise Ratio*, SNR) für Sprache im Hintergrundgeräusch. Dieser SNR wird Acceptable Noise Level (ANL) genannt. Er ist nach Nabelek *et al.* (1991) definiert als Differenz zwischen dem angenehmen Pegel der Sprache (engl. *Most Comfortable Level*, MCL) und dem maximalen Pegel eines Hintergrundgeräusches, den die Versuchsperson über längere Zeit beim Hören von Sprache akzeptieren würde. Dieser Pegel des Hintergrundgeräusches wird Background Noise Level (BNL) genannt. Der Begriff ANL ist missverständlich. Er beschreibt keinen Pegel sondern einen SNR, den die Versuchspersonen akzeptieren können. Personen mit einem niedrigen ANL akzeptieren dabei einen lauterer Pegel des Hintergrundgeräusches, während sie der Sprache zuhören, als Personen mit einem hohen ANL, die wenig Hintergrundgeräusch tolerieren.

Bisherige Untersuchungen zeigten u. a., dass der ANL unabhängig vom Grad des Hörverlustes, dem Alter, dem Geschlecht der Versuchsperson und deren Nationalität (Englisch, Koreanisch) und auch unabhängig davon ist, ob die Messungen mit oder ohne Hörgerät durchgeführt wurden (Nabelek *et al.*, 1991; Rogers *et al.*, 2003; Nabelek *et al.*, 2004, 2006; von Hapsburg und Bangh, 2006). Der ANL ist ebenso für die meisten Versuchspersonen nicht abhängig davon, ob monaural oder binaural mit Hörgeräten versorgt wurde (Freyaldenhoven *et al.*, 2006d). Außerdem zeigten Nabelek *et al.* (2004, 2006), dass der ANL nicht mit der Sprachverständlichkeit im Störgeräusch bei Messungen mit und ohne Hörgeräte zusammenhängt. Dagegen wurde festgestellt, dass der ANL von Hintergrundgeräuschen wie z. B. Stimmengewirr, stationärem Rauschen und Musik abhängig ist (Nabelek *et al.*, 1991; Smiley *et al.*, 2006; Gordon-Hickey und Moore, 2007). Außerdem zeigten Nabelek *et al.* (2006), dass der ANL zur Vorhersage der Hörgerätenutzung und damit der Zufriedenheit mit Hörgeräten eingesetzt werden kann, obwohl er nicht mit dem Sprachverstehen zusammenhängt. Freyaldenhoven *et al.* (2005, 2006e) und Mueller *et al.* (2006) wiesen mit dem ANLT den Nutzen von Hörgeräten mit Richtmikrofoncharakteristik, Zusatzbohrung in der Otoplastik, Verstärkungsausgleich der tiefen Frequenzen und digitalen Störgeräuschreduktionen nach.

Diese Studie legt den Schwerpunkt auf die Methode des ANLT. In acht unterschiedlichen Experimenten wurden die folgenden Fragestellungen untersucht:

- a) Welchen Einflusses hat die Schrittgröße von 5 bzw. 2 dB, 5 bzw. 1 dB und 2 bzw. 1 dB auf die Messwerte im ANLT?
- b) Welchen Einfluss hat eine auf- oder absteigende Instruktion im ANLT?
- c) Welchen Einfluss haben niedrige, mittlere und hohe Startpegel, bei denen die Signale zum ersten Mal im ANLT dargeboten werden?
- d) Wie groß ist die Test-Retest-Reliabilität und die Reliabilität über einen kurzen Zeitraum des ANLT und des Just Follow Conversation Test (JFCT), und können sich ANLT und JFCT gegenseitig beeinflussen?

- e) Welchen Einfluss hat die Methodik bei Wiederholungsmessungen?
- f) Welchen Einfluss nimmt eine SRT-Ermittlung mit dem Oldenburger Satztest auf den JFCT und den ANLT?
- g) Welchen Einfluss haben unterschiedliche Hintergrundgeräusche auf den ANL?
- h) Welchen Einfluss nimmt der Darbietungspegel in einer Trainingsphase, in der das Sprachmaterial geübt wird, auf den ANL?

Im Folgenden wird die allgemeine Methodik, die für alle Untersuchungen gleich ist, erklärt. Für jede Fragestellung getrennt wird danach die Zielsetzung, die Methode, die Versuchspersonen und die Ergebnisse beschrieben und diskutiert. Einige Experimente bauen aufeinander auf. Die Ergebnisse aller Experimente werden in den Schlussfolgerungen zusammengefasst.

3.3 Allgemeine Methodik

3.3.1 Apparatur

Für alle Messungen wurden die Signale von einem PC mit Mathworks Matlab (Version 7.3) verarbeitet, über eine Soundkarte (RME AD/DA-Interface ADI 2) an einen Kopfhörerverstärker (Tucker Davis Technologies HB7 Headphonedriver) übertragen und über einen Kopfhörer (Sennheiser HDA 200) in einer Hörkabine (Soundblocker, Audiometrikabine, Größe: B) wiedergegeben. Dieser Aufbau wurde freifeldentzerrt und nach DIN EN 60645-2 (1997) und DIN EN ISO 389-8 (2004) kalibriert. Anhang A enthält eine Beschreibung und die Kalibrierung des Messaufbaus.

3.3.2 Versuchspersonen

Vor Beginn der Untersuchungen wurde mit jeder Versuchsperson eine Tonaudiometrie durchgeführt. Die Versuchspersonen waren normalhörend. Sie hatten bei höchstens einer der Frequenzen 0,125, 0,25, 0,5, 1, 2, 4, 8 kHz einen Hörverlust von maximal 20 dB HL. Die Versuchspersonen waren z. T. Studenten und Mitarbeiter des Studienganges Hörtechnik und Audiologie in Oldenburg. Aber auch Personen, die sich auf öffentliche Ausgänge meldeten, nahmen an den Untersuchungen teil.

3.3.3 Testmaterial

Als Sprachsignal wurden die Sätze des Oldenburger Satztestes (OLSA) verwendet (Wagener *et al.*, 1999a,b,c; Wagener, 2003). Die Sätze haben eine einfache syntaktische Struktur und kaum Redundanz. Sie ergeben aufgrund ihrer Generierung nicht immer einen Sinn. Durch die Generierung und die gleiche syntaktische Struktur der Sätze wird verhindert, dass Versuchspersonen sie auswendig lernen können. 40 Testlisten können im OLSA verwendet werden, wobei eine Testliste aus 30 Sätzen besteht. Für eine kontinuierliche Wiedergabe wurden die Sätze in randomisierter Reihenfolge mit Pausen

zwischen 0,25 und 0,6 s dargeboten. Diese Werte liegen in einem Bereich, der von Camper und Veronis (2002) als mittlere Pausenlänge in fließend gesprochener Sprache bezeichnet wird. Das Rauschen des Oldenburger Satztestes wurde als Hintergrundgeräusch gewählt. Es hat das gleiche Spektrum wie die Sprache. Diese Signale wurden gewählt, um einen Vergleich der Ergebnisse des ANLT, des OLSA und des JFCT zu ermöglichen.

3.3.4 Durchführung des ANLT

Im ANLT wurden die Sätze des OLSA kontinuierlich dargeboten. Das weitere Vorgehen ist aus den Beschreibungen von Nabelek *et al.* (1991, 2006) entnommen. Die Versuchsperson stellte den Pegel der Sprache lauter als angenehm, danach leiser als angenehm und abschließend auf einen angenehmen Pegel (MCL), dem sie über eine längere Zeit zuhören würde, ein. Dieser Pegel der Sprache wurde konstant gehalten und in einem zweiten Schritt zusätzlich das Hintergrundgeräusch präsentiert. Der Pegel des Hintergrundgeräusches wurde dann durch die Versuchsperson verändert. Zuerst wählte sie eine Lautstärke, bei der sie die Sätze nicht mehr verstehen konnte, danach verringerte sie den Pegel, bis die Sätze deutlich zu verstehen waren. Zuletzt wählte sie die maximale Lautstärke des Hintergrundgeräusches, die sie über einen längeren Zeitraum tolerieren würde, ohne sich belästigt zu fühlen. Sie sollte den Sätzen zuhören können, ohne zu ermüden. Dabei sollte sie sich vorstellen, im Radio einer Dokumentation zu folgen und auf den Pegel eines Hintergrundgeräusches (z. B. eines Staubsaugers) direkt Einfluss nehmen zu können. In Anhang B.1 ist die Instruktion angefügt. Aus der Differenz zwischen MCL und BNL wurde der ANL berechnet. Zur Einstellung des Pegels wurde der Versuchsperson eine graphische Oberfläche (s. Anhang B.1) auf einem Bildschirm mit der aktuellen Instruktion präsentiert, und sie hatte die Möglichkeit, mit Schaltflächen die Lautstärke der Signale zu verändern. Mit Hilfe der Maus des PCs konnten die Schaltflächen ausgewählt werden. Sobald eine PegelEinstellung beendet war, bestätigte die Versuchsperson die Einstellung mit einem OK-Button, und die Instruktion für den nächsten Messschritt im ANLT wurde ihr präsentiert. Zur Untersuchung unterschiedlicher Merkmale wurden in der Regel je drei Bestimmungen des ANL durchgeführt. Dies ist die maximale Anzahl an Messwiederholungen, die in der Literatur beschrieben ist (Rogers *et al.*, 2003; Harkrider und Tampas, 2006; Mueller *et al.*, 2006; Franklin *et al.*, 2006).

3.3.5 Durchführung des OLSA

Der OLSA ist ein Sprachverständlichkeitstest im Störgeräusch, mit dem eine Sprachverständlichkeitsschwelle 50%-igen Satzverstehens ermittelt wurde (Wagener *et al.*, 1999a,b,c; Wagener, 2003). Dazu musste die Versuchsperson die im Störgeräusch dargebotenen Sätze wiederholen. Das Hintergrundgeräusch wurde, abhängig von den verstandenen Wörtern eines Satzes, in seinem Pegel verändert, bis ein SNR erreicht war, bei dem die Versuchsperson 50% der Sätze verstand. Dieser SNR wird Sprachverständlichkeitsschwelle (engl. *Speech Reception Threshold, SRT*) genannt. Es galt bei diesem

Vorgehen ein Satz bereits als falsch, wenn nur ein Wort falsch wiederholt wurde (engl. *Sentence-Scoring*) (Plomp und Mimpen, 1979; Nilsson *et al.*, 1994; Fredelake, 2006). Dies führte zu einer Erhöhung der Schwierigkeit für die Versuchsperson, die höhere SNR-Werte benötigte, um ein 50%-iges Satzverstehen zu erreichen. Damit entsprach die beschriebene Methode nicht dem von Wagener (2003) vorgestellten üblichen Verfahren, in dem die verstandenen Wörter jedes Satzes gezählt wurden und danach einen SNR-Wert für 50%-iges Wortverstehen eingeregelt wird (engl. *Word-Scoring*). Wagener *et al.* (1999b) empfehlen in einer Trainingsphase das Sprachmaterial zu üben, da es sonst innerhalb des Tests durch einen Lerneffekt zu Verbesserungen der Sprachverständlichkeit kommen kann. Deshalb wurde das Material mit Versuchspersonen, denen es nicht in ausreichendem Maße bekannt war, vor Beginn jedes Experimentes geübt. Dazu wurden zwei bis drei Testlisten mit 30 Sätzen zufällig ausgewählt und bei einem konstanten Pegel von 40 dB(A) ohne Hintergrundgeräusch den Versuchspersonen präsentiert. Sie hatten dann die Aufgabe, die Sätze zu wiederholen. Fehler wurden korrigiert, indem der richtige Satz erneut vorgesprochen wurde. Für die anschließenden Messungen erhielten die Versuchspersonen schriftliche und verbale Instruktionen. Sie sind in Anhang B dargestellt.

3.3.6 Durchführung des JFCT

Im JFC-Test (JFCT) wird ein subjektiv wahrgenommener SRT, der JFC-Wert, ermittelt (Hawkins und Stevens, 1950; Larsby und Arlinger, 1994; Borg *et al.*, 1998; Magnusson *et al.*, 2001; Neumann und Jensen, 2001; Latzel *et al.*, 2003). Der Versuchsperson wurden dazu die OLSA-Sätze kontinuierlich und zusammen mit dem OLSA-Rauschen dargeboten. Der Pegel des Hintergrundgeräusches konnte durch Bewegung einer Trackball-Maus verändert werden und SNR-Werte zwischen ± 15 dB SNR eingestellt werden. Aufgabe der Versuchsperson war es, den Pegel des Hintergrundgeräusches so zu wählen, dass sie die Sätze gerade eben noch verstehen konnte. Hatte sie diesen Pegel eingestellt, bestätigte sie ihn mit der Leertaste. Das Rauschen wurde nach Betätigung der Leertaste randomisiert um 5-10 dB angehoben oder abgesenkt, und die Versuchsperson stellte erneut ihre subjektive Sprachverständlichkeitsschwelle ein. Dieser Vorgang wurde je nach Untersuchung mehrmals wiederholt. Der Median der eingestellten SNR-Werte ergab den JFC-Wert.

3.4 Experiment A: Untersuchung der Schrittgröße

In der Literatur beschriebene Untersuchungen mit dem ANLT nutzten Schrittgrößen von 10, 2,5, 2 oder 1 dB zur Veränderung der Signalpegel (Nabelek *et al.*, 1991; Crowley und Nabelek, 1996; Rogers *et al.*, 2003; Nabelek *et al.*, 2004; Freyaldenhoven *et al.*, 2005; Harkrider und Tampas, 2006; Freyaldenhoven *et al.*, 2006c; Franklin *et al.*, 2006). In den neuesten Untersuchungen wird jedoch häufig eine Schrittgröße von 5 bzw. 2 dB für die Einpegelung verwendet (Harkrider und Smith, 2005; Nabelek *et al.*, 2006; von Hapsburg und Bangh, 2006; Freyaldenhoven *et al.*, 2006d; Mueller *et al.*, 2006). Dabei

findet die Schrittgröße von 5 dB bei der Einstellung der unteren und oberen Grenze der Sprache und des Rauschen Anwendung. Wird also Sprache zu laut und zu leise oder das Rauschen so laut, dass die Sprache nicht mehr verstanden und so leise, dass die Sprache deutlich verstanden wird, eingestellt, geschieht dies mit einer Schrittgröße von 5 dB. Zur Einstellung des angenehmen Pegels der Sprache und des maximal tolerierbaren Hintergrundgeräusches wird eine Schrittgröße von 2 dB verwendet. In diesem Experiment soll untersucht werden, ob diese Schritte ausreichend oder zu groß sind.

Methoden

Durchgeführt wurden ANL-Messungen nach in Kap. 3.3.4 beschriebenen Vorgehen mit den Schrittgrößen 5 bzw. 2 dB, 5 bzw. 1 dB und 2 bzw. 1 dB. Zuerst wurde in jeder Messung der MCL und danach der BNL bestimmt. Die Versuchspersonen führten drei ANL-Tests für jede Schrittgröße durch, so dass sie insgesamt neun Messungen absolvierten. Die Reihenfolge der Schrittgrößen wurde dabei randomisiert. Für jede Messung wurde außerdem die Anzahl der Pegelveränderungen dokumentiert. Dazu wurde gezählt, wie oft eine Versuchsperson die Schaltflächen zur Pegelveränderung verwendete.

Versuchspersonen

Am Experiment A nahmen elf Versuchspersonen (sechs Frauen, fünf Männer) im Alter zwischen 23 und 42 Jahren teil (mittleres Alter: 29,5 Jahre). Sie hatten geringe bzw. keine Erfahrung mit dem ANLT. Die Versuchspersonen waren normalhörend. Die Spanne ihres Hörvermögens ist in Anhang in Abb. C.1 dargestellt.

Ergebnisse

Um die Ergebnisse des ANLT auszuwerten, wurde aus den drei Messwiederholungen für jede Versuchsperson der Median errechnet und die Gesamtergebnisse aller Versuchspersonen in einem Boxplot dargestellt. Er beinhaltet Mediane, Quartile, Werte, die innerhalb der 1,5-fachen Quartilsspanne liegen und Ausreißer. Abb. 3.1 zeigt Boxplots aller Messwerte des ANLT bei unterschiedlichen Schrittgrößen. Die Messwerte sind bei allen Schrittgrößen sehr ähnlich. Der mediane MCL aller Messungen beträgt 50 dB(A) und wird in Kap. 3.5 als Startpegel verwendet. Die abgebildeten medianen ANL-Werte sind, auch wenn sie mit unterschiedlichen Schrittgrößen gemessen wurden, sehr ähnlich und liegen bei 10 dB SNR für die Schrittgrößen 5 bzw. 2 dB und 2 bzw. 1 dB und bei 9 dB SNR für die Schrittgröße 5 bzw. 1 dB SNR.

Mit einem Friedman-Test (Bortz *et al.*, 2000; Sachs, 2000) wurde die Signifikanz der ANL-Ergebnisse bei einem Signifikanzniveau von $\alpha = 0,05$ geprüft. Die ANL-Werte gemessen mit unterschiedlichen Schrittgrößen sind nicht signifikant unterschiedlich. Anhang D.1.1 zeigt die Ergebnisse des Friedman-Tests aller Messwerte des ANLT.

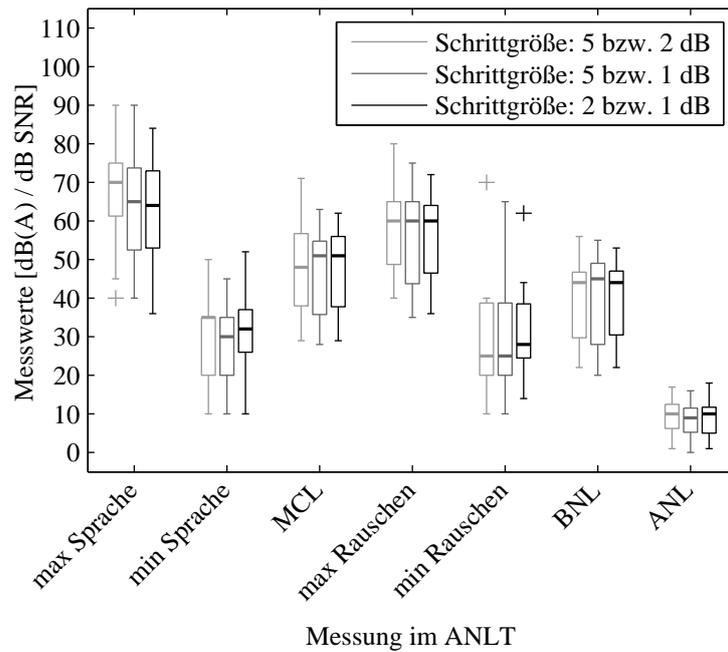


Abb. 3.1: Messwerte im ANLT, ermittelt mit unterschiedlichen Schrittgrößen. Die Ergebnisse maximale Sprache, minimale Sprache, MCL, maximales Rauschen, minimales Rauschen und BNL sind in dB(A) angegeben, die Werte des ANL in dB SNR.

Um einen Hinweis zur Streuung der Daten zu bekommen, wurden die Differenzen zwischen dem größten und dem kleinsten ANL-Wert für jede Schrittgröße und jede Versuchsperson ermittelt und in Abb. 3.2 dargestellt. Sie zeigt, dass mit einer Schrittgröße von 5 bzw. 2 dB im ANL im Median zwar eine Differenz von 4 dB SNR erreicht wurde, die Streuung dieser Daten aber sehr gering ist. In Anhang D.1.2 sind alle gemessenen ANL-Werte dargestellt.

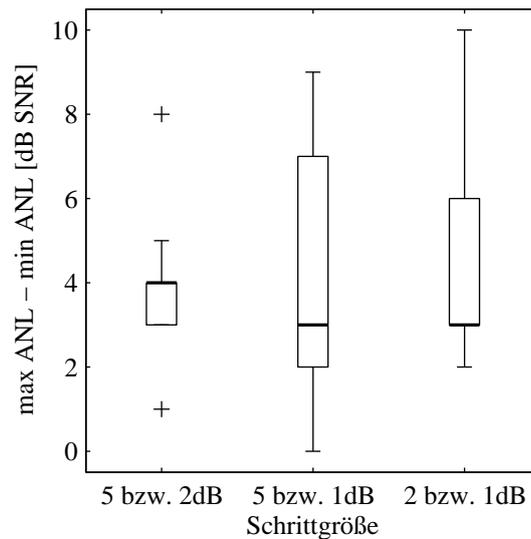


Abb. 3.2: Darstellung der Differenz zwischen maximalem und minimalem ANL-Wert bei unterschiedlichen Schrittgrößen.

Neben den ANL-Werten wurde die Anzahl der Pegelveränderungen zur Bestimmung eines Messpunktes, die von den Versuchspersonen vorgenommen wurden, ermittelt. Ihre Auswertung in Abb. 3.3 zeigt, dass mit einer Schrittgröße von 2 bzw. 1 dB und 5 bzw. 1 dB eine größere Anzahl an Veränderungen vorgenommen wurde.

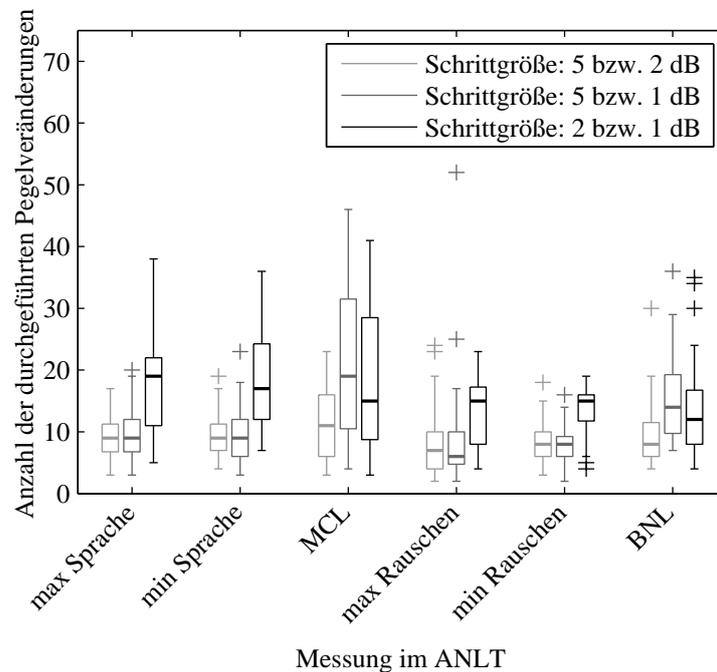


Abb. 3.3: Anzahl der Pegelveränderungen bei den einzelnen Messwerten im ANLT.

Diskussion

Im Experiment wurde der Einfluss der Schrittgrößen 5 bzw. 2 dB, 5 bzw. 1 dB und 2 bzw. 1 dB geprüft. D. h., die Einstellung der maximalen/minimalen Sprache und des maximalen/minimalen Rauschens wurde mit einer Schrittgröße von 5 oder 2 dB vorgenommen. Zur Einregelung des angenehmen Pegels der Sprache und des tolerierbaren Pegels des Hintergrundgeräusches wurden Schrittgrößen von 2 oder 1 dB verwendet. Die untersuchten Schrittgrößen führten nicht zu signifikanten Unterschieden des ANL. Mit der Schrittgröße 5 bzw. 2 dB wurden die geringste Anzahl an Pegelveränderungen vorgenommen (s. Abb. 3.3). Die Mehrzahl der Versuchspersonen bestätigte, dass sie eine Schrittgröße von 5 dB bei der Einstellung der Randbereiche („... Sprache lauter als angenehm ...“, „... Sprache leiser als angenehm ...“, „... Hintergrundgeräusch so laut, dass Sprache nicht mehr verstanden werden kann“, „... Hintergrundgeräusch so leise, dass Sprache sehr deutlich ist.“) bevorzugten. Sie beschrieben, dass die Randbereiche deutlich erreicht und überschritten werden konnten. Bei einer 2 dB-Schrittgröße wurden die Grenzen der Instruktionen undeutlich und als ein Bereich, der eingestellt werden konnte, wahrgenommen. Bei der Einstellung der „angenehmen Lautstärke“ („... Sprache auf eine sehr angenehme Lautstärke ...“, „... Hintergrundgeräusch auf eine maximal tolerierbare Lautstärke ...“) wurde eine Schrittgröße von 2 dB bevorzugt. Bei einer Schrittgröße von 1 dB war eine Veränderung des Pegels oft nicht wahrnehmbar. Die Versuchspersonen bevorzugten also eine Schrittgröße von 5 bzw. 2 dB, da die geforder-

ten Instruktionen damit schnell und präzise umgesetzt werden konnten. Auch bei den in der Literatur beschriebenen ANL-Tests werden meist Schrittgrößen von 5 bzw. 2 dB verwendet (Harkrider und Smith, 2005; Nabelek *et al.*, 2006; von Hapsburg und Bangh, 2006; Freyaldenhoven *et al.*, 2006d; Mueller *et al.*, 2006). Deshalb wurden sie auch als Schrittgrößen für die folgenden Tests beibehalten.

Schlussfolgerung des Experimentes A ist, dass mit der Schrittgröße 5 bzw. 2 dB der ANL schneller und präziser ermittelt werden kann als mit den Schrittgrößen 5 bzw. 1 dB und 2 bzw. 1 dB.

3.5 Experiment B: Untersuchung einer auf- oder absteigenden Instruktionen im ANLT

Das beschriebene Vorgehen beim ANLT beinhaltet eine vorgegebene Reihenfolge, in der zuerst ein maximaler danach ein minimaler und ein angenehmer Pegel der Sprache eingestellt wird (Nabelek *et al.*, 1991; Freyaldenhoven *et al.*, 2006d,e; Franklin *et al.*, 2006; Mueller *et al.*, 2006; von Hapsburg und Bangh, 2006; Harkrider und Tampas, 2006). Bei der Einstellung des Hintergrundgeräusches wird eine vergleichbare Reihenfolge verwendet, in der zuerst das Hintergrundgeräusch so laut eingestellt wird, dass Sprache nicht mehr verstanden werden kann bzw. so leise, dass Sprache sehr deutlich ist. Die Messwerte MCL und BNL werden somit immer von einem lauten Pegel kommend, d. h. in einer absteigenden Methode, eingestellt. Punch *et al.* (2004) untersuchten Methoden zur Ermittlung des MCL und beschrieben, dass bei Messungen mit einer absteigenden Methode 20-23 dB höhere MCL-Werte erreicht wurden als bei einer aufsteigenden Methode. Für den ANLT kann aus diesen Ergebnissen die folgende Hypothese abgeleitet werden: Wird diese Instruktion im ANLT in ihrer Reihenfolge verändert, d. h. werden erst minimale, dann maximale und anschließend angenehme Pegel eingestellt, dann liegen MCL und BNL möglicherweise höher als bei der herkömmlichen Reihenfolge, da sie von einem lauterem Pegel kommend bestimmt werden. Der ANL sollte jedoch bei beiden Instruktionsweisen gleiche Werte erreichen, da er als Differenz zwischen MCL und BNL definiert ist. Das hier durchgeführte Experiment prüft die Hypothese, dass, obwohl MCL und BNL durch eine Veränderung der Reihenfolge unterschiedlich sind, der ANL nicht beeinflusst wird.

Methode

Die Versuchspersonen führten Messungen mit zwei verschiedenen Instruktionen durch. Die Instruktion „Aufsteigend“ stellte die Aufgabe einen Pegel der Sprache einzustellen, der zu laut, dann zu leise und zum Schluss angenehm war und einen Pegel des Hintergrundgeräusches, der zuerst sehr laut war, so dass die Sprache nicht verstanden wurde, dann einen sehr leisen Pegel des Hintergrundgeräusches, so dass die Sprache sehr deutlich zu verstehen war und zum Schluss einen maximal tolerierbaren Pegel. Diese Instruktion wird im Folgenden „Aufsteigend“ genannt, da zur Einstellung des MCL und des BNL die Pegel erhöht werden sollten. Die Instruktion „Absteigend“ forderte die Versuchspersonen auf, die Sprache zuerst zu leise, dann zu laut und anschließend angenehm

einzustellen. Auch das Hintergrundgeräusch sollte zuerst sehr leise, so dass die Sprache sehr deutlich zu verstehen war und dann sehr laut, so dass die Sprache nicht verstanden wurde und zum Schluss auf einen maximal tolerierbaren Pegel eingestellt werden. Diese Instruktion wird „Absteigend“ genannt, da zur Einstellung des MCL und BNL die Pegel verringert werden müssen. Die verwendeten Instruktionen sind im Anhang B.2 aufgeführt. Die Versuchspersonen wurden in zwei Gruppen aufgeteilt und führten insgesamt sechs Bestimmungen des ANL durch, je drei Messungen mit den Instruktionen „Auf-“ und „Absteigend“. Eine Messung bestand aus der Bestimmung des ANL und des BNL. Es wurde darauf geachtet, dass die erste Gruppe mit der Instruktion „Aufsteigend“ und die zweite Gruppe mit der Instruktion „Absteigend“ begann. Der Startpegel, bei dem Sprache und Rauschen zum ersten Mal dargeboten wurden, wurde in diesem Experiment auf 50 dB(A) festgelegt. Dies entspricht dem Median der MCL-Werte aller Versuchspersonen in der Untersuchung der Schrittgröße (vgl. Kap. 3.4).

Versuchspersonen

An dieser Untersuchung nahmen elf normalhörende Versuchspersonen (zwei Frauen, neun Männer) im Alter zwischen 22 und 32 Jahren teil (mittleres Alter: 26,1 Jahre). Alle hatten keine Erfahrung mit dem ANLT. Im Anhang ist die Spanne des Hörvermögens der Versuchspersonen in Abb. C.2 dargestellt.

Ergebnisse

Aus den drei Messwiederholungen jeder Versuchsperson wurden Mediane ermittelt. Die Ergebnisse aller Versuchspersonen wurden in einem Boxplot dargestellt und ihre Signifikanz mit einem Wilcoxon Vorzeichen-Rang-Test (Bortz *et al.*, 2000; Sachs, 2000) geprüft. In Anhang D.2.1 sind die Ergebnisse des ANLT in der gemessenen Reihenfolge dargestellt. Abb. 3.4 zeigt einen Boxplot der einzelnen Messwerte des ANLT, die mit den Instruktionen „Auf-“ und „Absteigend“ erreicht wurden. Die Mediane und ihre Streuung sind den im Test gewählten Instruktionen bzw. Messschritten zugeordnet. Vor allem für den MCL und den BNL ist zu erkennen, dass die Mediane höher liegen, wenn die Instruktion „Absteigend“ verwendet und damit zuvor der maximale Pegel eingestellt wurde. Die Medianwerte des ANL liegen bei 3 dB SNR mit dem Instruktion „Auf-“ und „Absteigend“. Nach einem Wilcoxon Vorzeichen-Rang-Test mit einem Signifikanzniveau von $\alpha = 0,05$ sind die ANL-Werte nicht signifikant unterschiedlich. MCL- und BNL-Werte, gemessen mit der auf- und absteigenden Instruktion, unterscheiden sich jedoch signifikant (MCL: $p = 0,003$, BNL: $p = 0,007$). Die Irrtumswahrscheinlichkeiten aller Messwerte des ANLT ermittelt dem Wilcoxon Vorzeichen-Rang-Test sind in Anhang D.2.2 dargestellt.

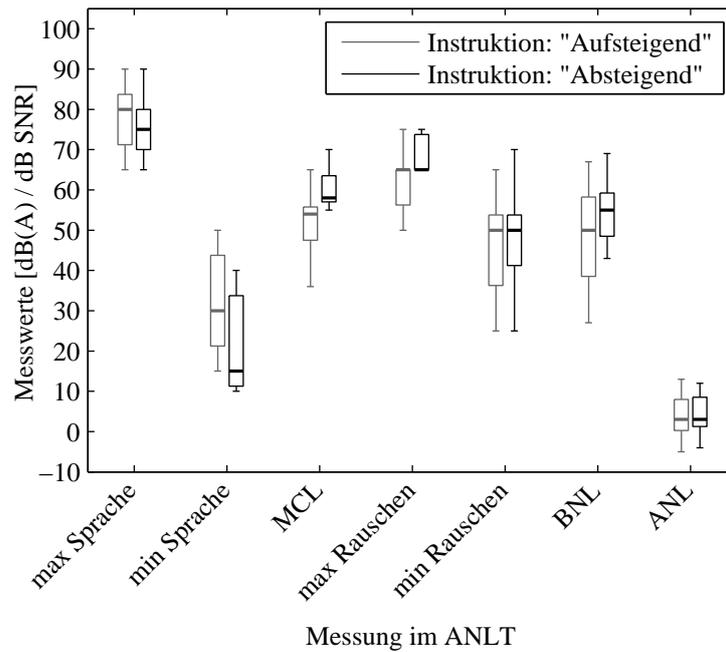


Abb. 3.4: Messwerte im ANLT, gemessen mit auf- oder absteigender Instruktion. Die gemessenen Werte wurden den im Test verwendeten Instruktionen zugeordnet. Die Ergebnisse maximale Sprache, minimale Sprache, MCL, maximales Rauschen, minimales Rauschen und BNL sind in dB(A) und die Werte des ANL in dB SNR angegeben.

Diskussion

Die Ergebnisse zeigen, dass die Reihenfolge der Instruktion zu unterschiedlichen Werten der minimalen Sprache, des MCL und des BNL führt (s. Abb. 3.4). Ursache ist vermutlich, wie schon von Punch *et al.* (2004) beschrieben, die Richtung der Pegelveränderung. Bei einer aufsteigenden Methode werden niedrigere Werte erreicht als bei einer absteigenden Methode. Wahrscheinlich nehmen die Versuchspersonen in einer absteigenden Methode bei einem höheren Pegel an, eher den gewünschten Wert erreicht zu haben, als bei einer aufsteigenden Methode. Einen Einfluss auf das Endergebnis, den ANL, hat die Änderung der Messreihenfolge jedoch nicht. Der ANLT, wie er in der Literatur beschrieben wird (Nabelek *et al.*, 1991; Freyaldenhoven *et al.*, 2006d,e; Franklin *et al.*, 2006; Mueller *et al.*, 2006; von Hapsburg und Bangh, 2006; Harkrider und Tampas, 2006), sieht vor, zuerst einen Pegel der Sprache zu ermitteln, der zu laut, danach zu leise und zum Schluss angenehm ist und einen Pegel des Hintergrundgeräusches, der sehr laut ist, so dass die Sprache nicht zu verstehen ist, dann einen Pegel, der sehr leise ist, so dass die Sprache sehr deutlich zu verstehen ist, und abschließend einen maximal tolerierbaren Pegel. Aufgrund dieser Ergebnisse und der verwendeten Methode in den o. g. Studien wird an dieser Reihenfolge auch in den folgenden Experimenten festgehalten.

Fazit des Experimentes B ist, dass die Veränderung der Reihenfolge der Instruktion im ANLT zu unterschiedlichen Messergebnissen des MCL und BNL führt. Der ANL, die Differenz des MCL und BNL, wird jedoch nicht beeinflusst.

3.6 Experiment C: Untersuchung des Startpegels

Schick (1979) wies darauf hin, dass in subjektiven Messungen der Startpegel, bei dem die Signale in einer Messung das erste Mal dargeboten werden, Einfluss auf die Endergebnisse nehmen kann. Er kann von der Versuchsperson als Anker wahrgenommen werden, so dass sie während der Messung Pegel wählt, die nicht weit von dem Startpegel entfernt liegen. Für den originalen ANLT wird ein Startpegel von 30 dB HL vorgeschlagen (Freyaldenhoven *et al.*, 2005; Nabelek *et al.*, 2006; Freyaldenhoven *et al.*, 2006d,c; Gordon-Hickey und Moore, 2007). Andere Studien verwenden Startpegel von z. B. 0 dB HL (Rogers *et al.*, 2003; Harkrider und Tampas, 2006), 10 dB HL (Harkrider und Smith, 2005), 40 dB SPL (Mueller *et al.*, 2006) oder 70 dB SPL (von Hapsburg und Bangh, 2006). In diesem Experiment soll untersucht werden, ob hohe, mittlere und niedrige Startpegel für die Sprache und das Hintergrundgeräusch Einfluss auf die Ergebnisse des ANLT haben.

Methoden

Durchgeführt wurde der ANLT mit geringen, mittleren und hohen Startpegeln von 30, 50 und 70 dB(A). Jeder Startpegel wurde den Versuchspersonen dreimal dargeboten, so dass sie insgesamt neun Messungen absolvierten. Jede Messung begann mit der Einpegelung der Sprache und wurde fortgesetzt mit der Einstellung des Hintergrundgeräusches. Die Reihenfolge der dargebotenen Startpegel wurde randomisiert.

Versuchspersonen

An dieser Untersuchung nahmen zwölf normalhörende Versuchspersonen (sechs Frauen, sechs Männer) teil. Sie hatten alle bereits Erfahrung aus vorhergehenden Messungen mit dem ANLT und waren zwischen 21 und 41 Jahren alt (mittleres Alter: 27,2 Jahre). Die Spanne des Hörvermögens der Versuchspersonen ist in Abb. C.3 im Anhang dargestellt.

Ergebnisse

Zur Auswertung der Ergebnisse wurde der Median der drei Messwiederholungen jeder Versuchsperson ermittelt. Abb. 3.5 zeigt einen Boxplot der Messwerte im ANLT mit den Startpegeln 30, 50 und 70 dB(A). Alle ermittelten Werte zeigen keine großen Unterschiede. Der mediane ANL beträgt bei einem Startpegel von 30 dB(A) 10 dB SNR. Bei einem Startpegel von 50 dB(A) wird ein medianer ANL von 9,5 dB SNR und mit 70 dB(A) ein medianer ANL von 11 dB SNR ermittelt. Damit zeigen Messungen bei einem Startpegel von 70 dB(A) einen etwas höheren Median als die Messungen mit Startpegeln bei 50 oder 30 dB(A). Die Signifikanz wurde bei einem Signifikanzniveau von $\alpha = 0,05$ mit einem Friedman-Test geprüft und signifikante Unterschiede der ANL-Werte ermittelt ($\chi^2_{2;22;0,05} = 7,143, p = 0,028$). Nach einem paarweisen Vergleich mit einem Wilcoxon Vorzeichen-Rang-Test zeigte sich, dass sich die ANL-Werte, ermittelt mit einem Startpegel von 50 und 70 dB(A), auf einem Niveau von $\alpha = 0,05$ signifikant

unterscheiden ($p = 0,031$). Die Differenz dieser ANL-Mediane beträgt 1,5 dB SNR. Wird eine Bonferroni-Adjustierung des Signifikanzniveaus ($\alpha = \frac{0,05}{3} = 0,017$) vorgenommen, ist der Unterschied nicht mehr signifikant. Die Ergebnisse eines Friedman-Tests aller Messwerte des ANLT sind in Anhang D.3 dargestellt.

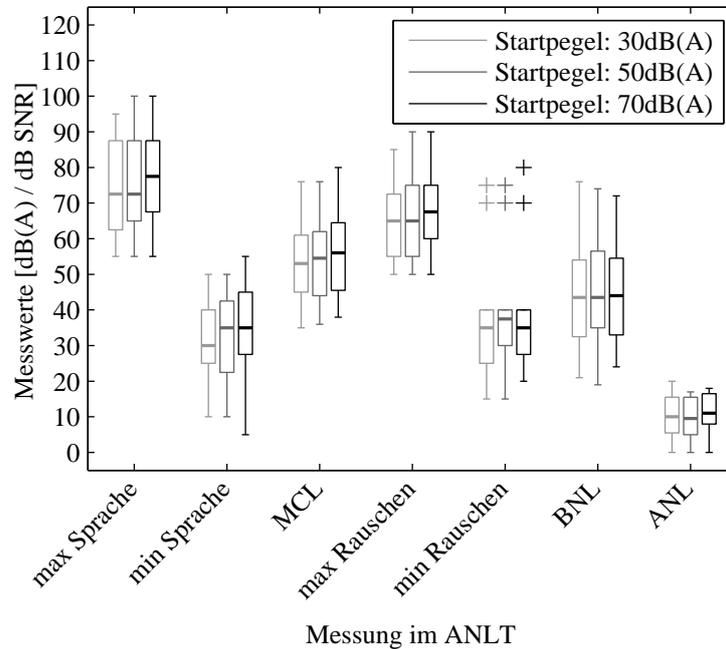


Abb. 3.5: Messwerte im ANLT, gemessen bei unterschiedlichen Startpegeln. Die Ergebnisse maximale Sprache, minimale Sprache, MCL, maximales Rauschen, minimales Rauschen und BNL sind in dB(A), die Werte des ANL in dB SNR angegeben.

Diskussion

Die Ergebnisse zeigen, dass ein Startpegel von 70 dB(A) zu tendenziell höheren ANL-Werten führen kann (s. Abb. 3.5). Die Differenz der Mediane der ANL-Werte ist mit ca. 1,5 dB SNR jedoch sehr gering. Für die Startpegel von 50 dB(A) oder 30 dB(A) konnte kein signifikanter Unterschied ermittelt werden. Diese Ergebnisse können genutzt werden, wenn mehrere Messwiederholungen vorgenommen werden sollen. Verschiedene Startpegel können in diesem Fall dazu genutzt werden, dem Zählen der vorgenommenen Pegelveränderungen durch die Versuchsperson und der damit verbundenen Routine vorzubeugen. Ist eine Randomisierung des Startpegels nicht notwendig, da z. B. Messungen mit unterschiedlichen Hintergrundgeräuschen durchgeführt werden, kann der Startpegel von 30 dB(A) beibehalten werden. Diese Untersuchung sollte mit einer größeren Anzahl an Versuchspersonen wiederholt werden, um den Einfluss lauter Startpegel zu prüfen.

Ergebnis des Experiment C ist, dass niedrige und mittlere Startpegel von 30 bzw. 50 dB(A) keinen Einfluss auf den ANL haben. Hohe Startpegel von 70 dB(A) weisen die Tendenz auf, den ANL zu höheren Pegeln zu verschieben. Bei Wiederholungsmessungen ist also eine Veränderung des Startpegels möglich, um routiniertem Antwortverhalten vorzubeugen.

3.7 Experiment D: Test-Retest-Reliabilität und Reliabilität über einen kurzen Zeitraum

Nachdem in den zuvor durchgeführten Experimenten die Schrittgröße, die Instruktion und der Startpegel des ANLT untersucht wurden, soll nun die Test-Retest-Reliabilität (zu zwei unterschiedlichen Terminen) und die Reliabilität über einen kurzen Zeitraum (innerhalb eines Messtermins) geprüft werden. Dazu wird der Test an zwei unterschiedlichen Terminen mehrmals zusammen mit dem JFCT durchgeführt. Außerdem soll untersucht werden, ob sich ANLT und JFCT gegenseitig beeinflussen.

Untersuchungen zur Test-Retest-Reliabilität des ANLT wurden von Nabelek *et al.* (2004) mit Messungen von einer Woche Abstand und von Freyaldenhoven *et al.* (2006c) mit Messungen von ein bis zwei Monaten Abstand vorgenommen. In beiden Studien wurden hochsignifikante Korrelationen von 0,79-0,81 (Freyaldenhoven *et al.*, 2006c) und 0,79-0,85 (Nabelek *et al.*, 2004) zwischen den ANL-Werten zu unterschiedlichen Zeitpunkten erreicht. Latzel *et al.* (2003) entwickelten die HörTech Einpegelung der Individuellen Satzverständlichkeitsschwelle (HEISS), deren Methode dem JFCT sehr ähnlich ist. Abweichend vom JFC-Verfahren werden beim HEISS-Test Sätze des Göttinger Satztestes (Kollmeier und Wesselkamp, 1997) verwendet. Dieses Verfahren zeigte eine Korrelation von 0,69 bei Messungen, die in zwei unterschiedlichen Terminen durchgeführt wurden. Latzel *et al.* (2003) erklären diesen geringen Wert durch eine Veränderung des JFC-Kriteriums, also der individuellen internen Referenz. Sie folgerten daraus, dass die Ergebnisse keinen Zusammenhang zeigen und ein Vergleich von Ergebnissen, gemessen in unterschiedlichen Terminen, unzulässig ist.

Methode

ANLT und JFCT wurden an zwei verschiedenen Terminen mit ca. einer Woche Abstand durchgeführt. In einem Cross-Over-Design wurde einer ersten Gruppe (Gruppe 1) im ersten Termin (Termin A) zuerst der ANLT zwölfmal und anschließend der JFCT zwölfmal präsentiert. Diese Messreihenfolge wird im Folgenden mit ANLT-JFCT gekennzeichnet. Im zweiten Termin (Termin B) wurde mit dieser Gruppe zuerst der JFCT und darauf folgend der ANLT ausgeführt. Diese Messreihenfolge wird JFCT-ANLT genannt. Mit einer zweiten Gruppe (Gruppe 2) wurden die Messwerte in umgekehrter Reihenfolge ermittelt (Termin A: JFCT-ANLT, Termin B: ANLT-JFCT). Der Startpegel wurde beim ANLT zwischen den Werten 20, 30 und 40 dB(A) randomisiert. Diese Pegel wurden gewählt, damit den Versuchspersonen das Zählen der vorgenommenen Pegelveränderungen nicht möglich war. Damit schwankten die Pegel um den in der Literatur häufig verwendeten Startpegel von 30 dB (s. Kap. 3.6).

Bei der Durchführung der JFC-Tests der Gruppe 1 wurde der Pegel der Sprache auf den im ANLT ermittelten MCL eingestellt. Um einen MCL für die Gruppe 2 zu ermitteln, wurde jede Versuchspersonen in einer Vormessung dreimal gebeten, einen angenehmen Pegel der Sprache einzustellen. Der Startpegel der Sprache betrug hierbei 30 dB(A) und konnte in 2 dB-Schritten verändert werden. Der Median der drei Messungen wurde als Pegel der Sprache im JFCT verwendet.

Da von Wagener *et al.* (1999b) ein Trainingseffekt im OLSA bei der Ermittlung eines SRT beschrieben wurde, wurde das OLSA-Sprachmaterial für die Ermittlung der subjektiven Sprachverständlichkeitsschwelle im JFCT geübt. Das Training wurde vorgenommen, wenn der JFCT im Termin A bzw. B als erste Messung durchgeführt wurde. Wurde der JFCT als zweite Messung durchgeführt, wurde auf eine Trainingsphase verzichtet, da die Versuchspersonen das Sprachmaterial im ANLT bereits gehört hatten. Wurde der ANLT im Termin A bzw. B als erste Messung durchgeführt, so wurde nur mit Versuchspersonen, denen das Sprachmaterial nicht in ausreichendem Maße bekannt war, vor Beginn jedes Experimentes geübt. Den Versuchspersonen wurden zum Training eine bis drei Listen bei einem festen Pegel von 40 dB(A) ohne Hintergrundgeräusch präsentiert, und sie wiederholten die verstandenen Wörter. Fehler wurden korrigiert, indem der Satz erneut wiederholt wurde.

Versuchspersonen

An dieser Untersuchung nahmen zwölf normalhörende Versuchspersonen (fünf Frauen, sieben Männer) in einem Alter zwischen 20 und 37 Jahren teil (mittleres Alter: 25,7 Jahre). Sie hatten alle keine Erfahrung mit dem ANLT. Die Spanne ihres Hörvermögens ist in Anhang C.4 dargestellt.

Ergebnisse

Ergebnisse des ANLT Zur Darstellung der Reliabilität über einen kurzen Zeitraum wurden in Abb. 3.6 (a) die MCL-Werte, (b) die BNL-Werte und (c) die ANL-Werte, die bei bis zu zwölfmaliger Wiederholung ermittelt wurden, in Boxplots zusammengefasst. Getrennt dargestellt sind die Ergebnisse der Termine A und B. Die MCL-Werte des Termins B liegen höher als die des Termins A (s. Abb. 3.6a). Außerdem zeigen alle MCL-Werte die Tendenz, bei einer höheren Anzahl an Messwiederholungen zu höheren Pegeln verschoben zu sein. Die Prüfung dieses Ergebnisses mit einem Friedman-Test ergab einen signifikanten Unterschied zwischen den MCL-Werten der Termine A und B auf einem Signifikanzniveau von $\alpha = 0,05$ ($\chi^2_{(1;246;0,05)} = 82,498$, $p = 0$). Ein deutlicher Unterschied wie bei den MCL-Werten der Termine A und B ist zwischen den BNL-Werten der unterschiedlichen Termine nicht zu erkennen (s. Abb. 3.6b). Nach einem Friedman-Test sind sie jedoch signifikant unterschiedlich ($\chi^2_{(1;246;0,05)} = 51,723$, $p \approx 0$). Abb. 3.6c zeigt eine Erhöhung der ANL-Werte, je mehr Messungen durchgeführt wurden. Dieser Trend ist bei den Werten der Termine A und B zu erkennen. Ein Friedman-Test ergab keine signifikanten Unterschiede zwischen den ANL-Werten der beiden Termine ($\chi^2_{(1;246;0,05)} = 3,269$, $p = 0,071$).

Zur Untersuchung der Test-Retest-Reliabilität wurde für jede Versuchsperson der Median der zwölf Messwiederholungen der Termine A und B ermittelt. Für die MCL-, BNL- und ANL-Werte der Termine A und B wurde der Spearmansche Korrelationskoeffizient r_{sp} und seine Irrtumswahrscheinlichkeit p mit einem Permutationstest (Sachs, 2000) berechnet. Der Permutationstest prüft die Hypothese, dass die Korrelation signifikant unterschiedlich von $r_{sp} = 0$ ist. Die Korrelationen r_{sp} liegen für den MCL bei 0,83, den

BNL bei 0,86 und den ANL bei 0,87 und sind signifikant auf einem Niveau von $\alpha = 0,01$. Im Anhang D.4.1 sind die Zusammenhänge als Scatterplot dargestellt.

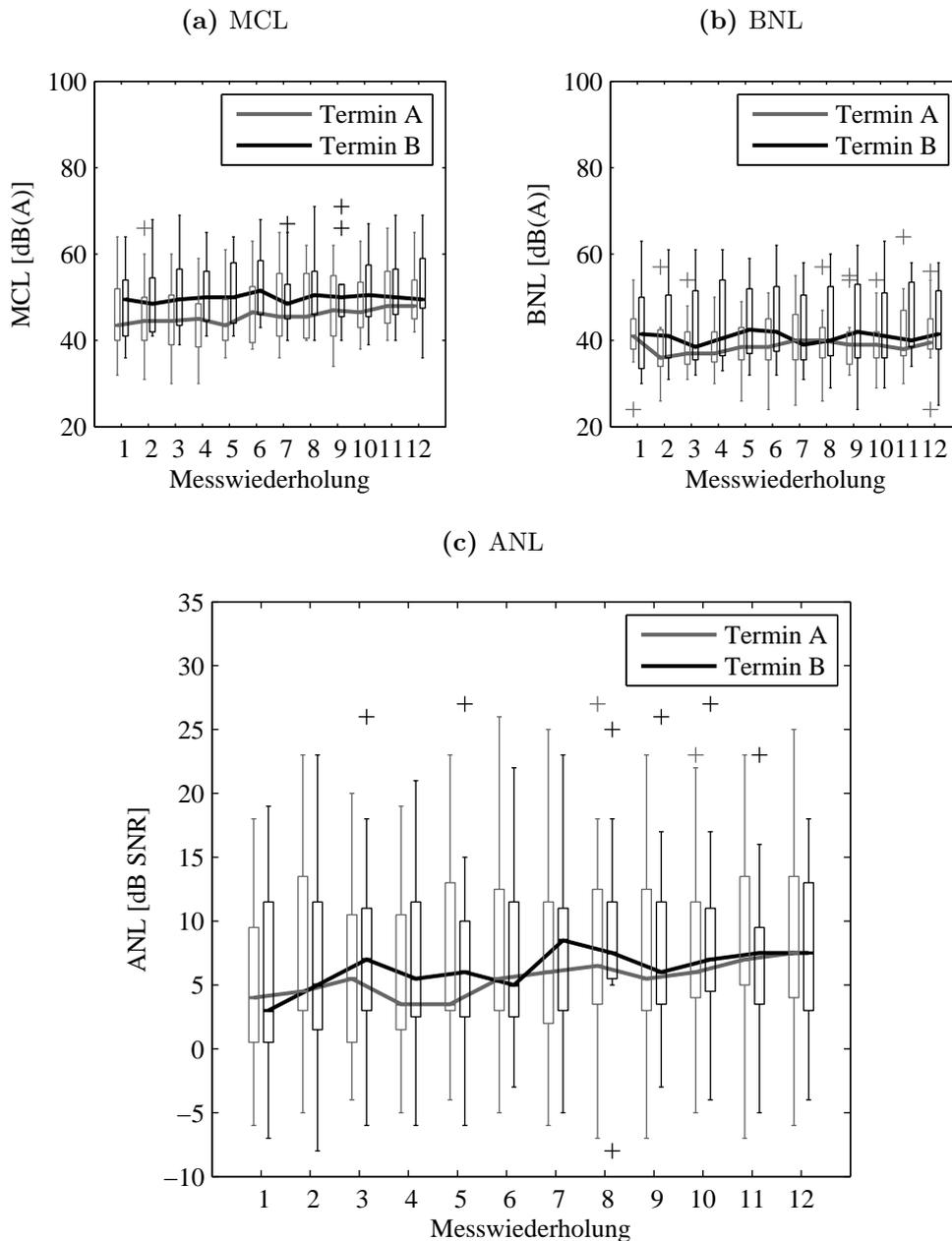


Abb. 3.6: (a) MCL-, (b) BNL- und (c) ANL- Werte, gemessen bei bis zu zwölf Messwiederholungen in den Terminen A und B.

Der Interquartilsabstand wurde zur Untersuchung der Streuung des ANL berechnet. Die mediane Streuung des Termins A beträgt 2,5 dB SNR, die des Termins B 3 dB SNR. Ihre Maxima und Minima liegen bei 5,5 dB SNR und 1,5 dB SNR (Termin A) bzw. 6,5 dB SNR und 1,5 dB SNR (Termin B). Im Anhang D.4.2 sind die inter- und die intraindividuelle Streuung jeder Versuchsperson dargestellt.

Um die Anzahl der nötigen Wiederholungen der Messung zu ermitteln, wurde eine Kumulierung der Ergebnisse vorgenommen. Für die Abb. 3.7 wurde der Median der ersten ANL-Werte jeder Versuchsperson ermittelt. Kumuliert wurde die auf der Abszisse do-

kumentierte Anzahl der Werte. Ist die Anzahl der gemittelten Messungen z. B. drei, so wurde für jede Versuchsperson der Median der ANL-Werte, die sie bei den ersten drei Messungen eingestellt hatte, gebildet. Auch die kumulierten Werte zeigen einen Anstieg, je mehr Datenwerte berücksichtigt werden. Außerdem wurden die kumulierten Werte mit Hilfe des Friedman-Tests auf signifikante Unterschiede bei einem Signifikanzniveau von $\alpha = 0,05$ untersucht. Die signifikanten Ergebnisse sind in Anhang D.4.3 aufgelistet. Erst bei einem Vergleich der Kumulierung von neun Werten mit zwölf kumulierten Werten entstehen keine signifikanten Unterschiede mehr. Das bedeutet, dass eine Kumulierung über die ersten neun oder ersten zwölf ANL-Werte zu keinem Unterschied in den Ergebnissen führt und eine Messung von neun Werten ausreichend wäre, wenn sie mit einer Messung von zwölf Werten verglichen wird.

Eine von der interindividuellen Streuung befreite Betrachtung der Ergebnisse der Abb. 3.6c und 3.7 ist in Anhang D.4.4 aufgeführt. Da die in Kap. 3.6 untersuchten Startpegel nicht mit den hier verwendeten Startpegeln übereinstimmen, wurde die Abhängigkeit der ANL-Werte vom Startpegel mit einem Friedman-Test geprüft. Signifikante Unterschiede durch den Startpegel konnten nicht festgestellt werden ($\chi^2_{(2;66;0,05)} = 0,149$, $p = 0,928$).

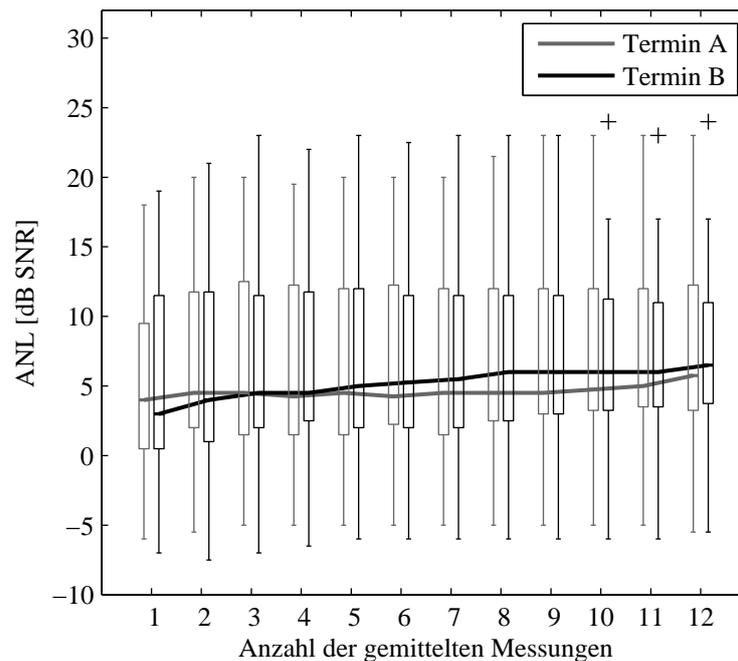


Abb. 3.7: Kumulierte ANL-Werte jeder Versuchsperson, gemessen in den Terminen A und B. Darin sind Medianwerte jeder Versuchsperson zusammengefasst. Für jede Versuchsperson wurden aus der auf der Abszisse angegebene Anzahl der ANL-Werte der Median ermittelt. Ist die Anzahl der gemittelten Messungen z. B. drei, so wurde für jede Versuchsperson der Median der Werte, die sie bei den ersten drei Messungen eingestellt hatte, gebildet.

Ergebnisse des JFCT Zur Analyse der Ergebnisse der JFCT wurden ebenso wie beim ANLT die Einzelwerte, Interquartilsabstände und kumulierte JFC-Werte berechnet. Zur Darstellung der Reliabilität über einen kurzen Zeitraum zeigt Abb. 3.8 Boxplots der ermittelten JFC-Werte bei bis zu zwölf Messwiederholungen der Termine A und B. Es ist ein geringer Abfall der Werte zwischen den ersten und den folgenden Messwerten vor allem für die JFC-Werte des Termins B zu erkennen, je mehr Messungen durchgeführt wurden. Die Prüfung der Ergebnisse der zwei Messzeitpunkte mit einem Friedman-Test ergab einen signifikanten Unterschied auf einem Niveau von $\alpha = 0,05$ ($\chi^2_{(1;246;0,05)} = 43,521$, $p \approx 0$). Zur Untersuchung der Test-Retest-Reliabilität wurde aus den Werten der zwölf Messwiederholungen für jede Versuchsperson der Median berechnet und der Spearmansche Korrelationskoeffizient r_{sp} bestimmt. Die Korrelation der JFC-Werte der Termine A und B von $r_{sp} = 0,692$ ist signifikant ($p = 0,016$, $\alpha = 0,05$). Anhang D.4.5 zeigt die Gegenüberstellung der JFC-Werte.

Zur Untersuchung der Streuung der JFC-Werte wurde auch hier der Interquartilsabstand verwendet. Die mediane Streuung beträgt 1,3 dB SNR (Termin A) und 1 dB SNR (Termin B). Ihre Maxima liegen bei 1,8 dB SNR (Termin A) und 1,4 dB SNR (Termin B) und die Minima bei 0,9 dB SNR (Termin A) bzw. 0,7 dB SNR (Termin B). Im Anhang D.4.6 ist die inter- und intraindividuelle Streuung dargestellt.

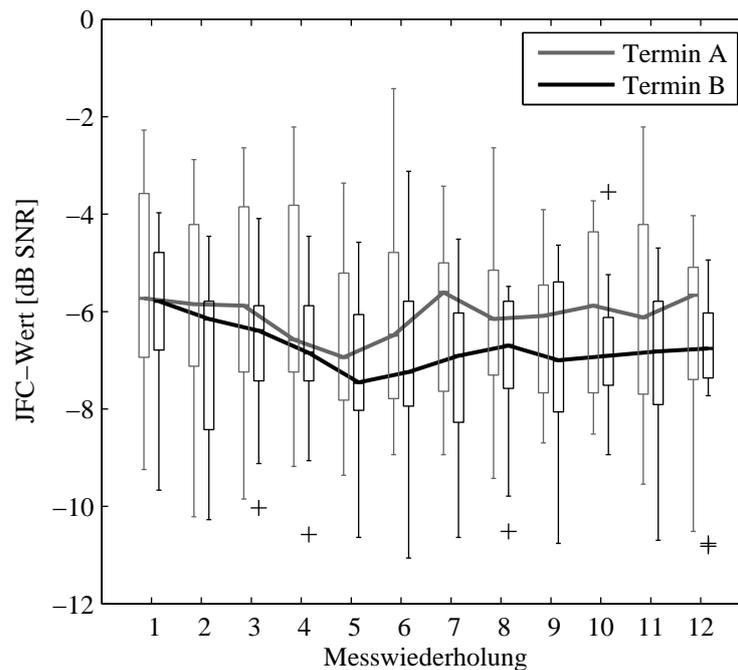


Abb. 3.8: JFC-Werte, gemessen bei bis zu zwölf Messwiederholungen in den Terminen A und B.

Für Abb. 3.9 wurde der Median der JFC-Werte errechnet. Kumuliert wurde dazu die auf der Abszisse angegebene Anzahl der Werte. Ist die Anzahl der gemittelten Messungen z. B. fünf, so wurde für jede Versuchsperson der Median der JFC-Werte, die sie bei den ersten fünf JFC-Messungen eingestellt hatte, gebildet. Alle Kumulierungen wurden auf signifikante Unterschiede mit dem Friedman-Test mit einem Signifikanzniveau von $\alpha = 0,05$ untersucht. Zwischen der Kumulierung der ersten vier und der ersten acht Wer-

te kann ein signifikanter Unterschied festgestellt werden ($\chi^2_{1;24;0,05} = 4,19, p = 0,041$). Vergleiche mit einer höheren Anzahl an Kumulierungen zeigten kein signifikantes Ergebnis im Friedmantest. Eine Darstellung aller signifikanten Unterschiede der kumulierten JFC-Werte findet sich in Anhang D.4.7 und eine von der interindividuellen Streuung unabhängige Betrachtung der JFC-Werte ist in Anhang D.4.8 dargestellt.

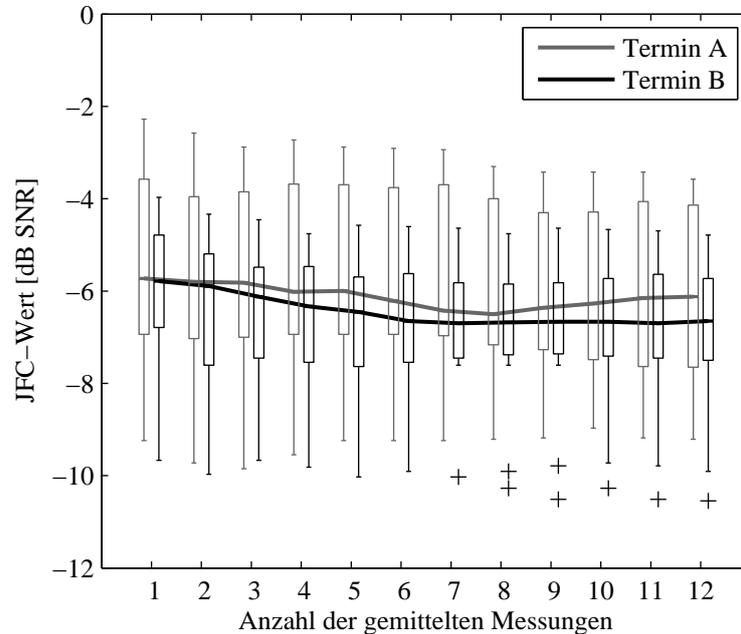


Abb. 3.9: Kumulierte JFC-Werte jeder Versuchsperson, gemessen in den Terminen A und B. Für jede Versuchsperson wurden aus der auf der Abszisse angegebenen Anzahl der JFC-Werte der Median ermittelt. Ist die Anzahl der gemittelten Messungen z. B. drei, so wurde für jede Versuchsperson der Median der Werte, die sie bei den ersten drei Messungen eingestellt hatte, gebildet.

Abhängigkeit von ANLT und JFCT Um den Einfluss des JFCT auf den ANLT zu prüfen, wurden die Ergebnisse aller Versuchspersonen getrennt nach der Messreihenfolge betrachtet. Zum Vergleich alle Ergebnisse miteinander wurden in Abb. 3.10 die Mediane der bis zu zwölf Messwiederholungen jeder Testsituation aufgetragen. Die Messungen, die mit der gleichen Reihenfolge durchgeführt wurden, sind mit einem gleichen Symbol dargestellt. Messungen, die zum gleichen Termin durchgeführt wurden, haben die gleiche Farbe, und Messungen der gleichen Versuchspersonengruppe haben die gleiche Verbindungslinie. Die medianen ANL-Werte der Gruppe 2, die im Termin B in der Reihenfolge ANLT-JFCT gemessen wurden, weichen von den anderen ANL-Werten deutlich zu höheren Werten ab.

Um die Signifikanz der Unterschiede der Werte zu prüfen, wurde zuerst für jede Versuchsperson der Median aus ihren zwölf Messergebnissen ermittelt. Mit dem Mann-Whitney-U-Test (U) und dem Kolmogorov-Smirnov-Test (KS, Bortz *et al.* (2000); Sachs (2000)) wurden dann die Werte des Termins A der Reihenfolge ANLT-JFCT und JFCT-ANLT auf signifikante Unterschiede der zentralen Tendenz und der gesamten Verteilung auf einem Signifikanzniveau von $\alpha = 0,05$ untersucht. Beide Tests zeigen keinen signifikanten Unterschied ($p_U = 0,699, p_{KS} = 0,81$). Die Irrtumswahrscheinlichkeiten p werden

höher, wenn auch die Werte des Termins B mit berücksichtigt werden ($p_U = 0,707$, $p_{KS} = 0,991$). Werden nur die Ergebnisse des Termins B mit den Signifikanztests verglichen, so sind auch sie nicht signifikant unterschiedlich ($p_U = 0,286$, $p_{KS} = 0,81$).

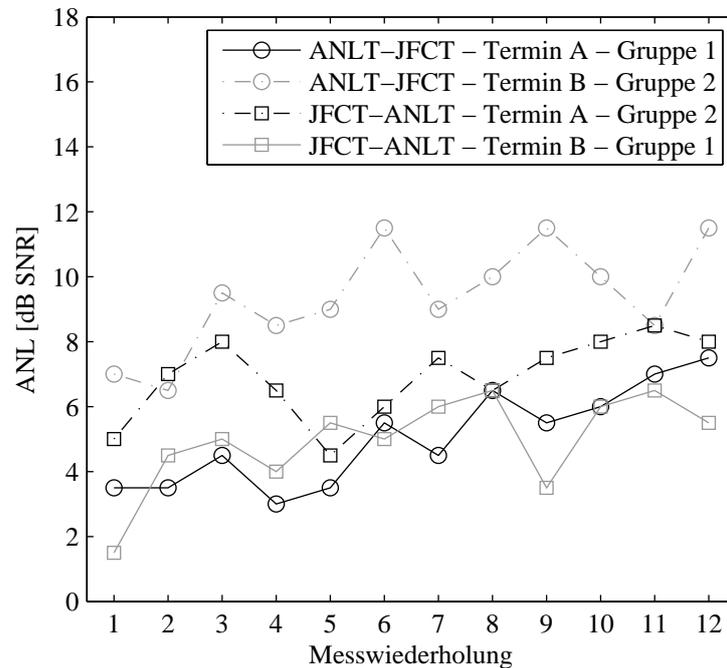


Abb. 3.10: Mediane der ANL-Werte. Dargestellt sind die Medianwerte, die für jede ANL-Messung mit den unterschiedlichen Parametern Reihenfolge, Termin und Versuchspersonengruppe ermittelt wurden.

Auf gleiche Weise wie zuvor die ANL-Werte wurden die Ergebnisse der JFC-Werte ausgewertet. Zur Verdeutlichung sind die medianen JFC-Werte der bis zu zwölf Messwiederholungen jeder Testsituation zusammen in Abb. 3.11 dargestellt. Auch in dieser Abbildung sind gleiche Messreihenfolgen durch die Symbole, gleiche Termine durch die Farbe und gleiche Gruppen durch die Linien gekennzeichnet. Die JFC-Werte der Messreihenfolge ANLT-JFCT des Termin A (Gruppe 1) weichen von den anderen zu höheren Werten ab.

Zur Untersuchung signifikanter Unterschiede wurde wiederum aus den zwölf Messwiederholungen der JFC-Werte für jede Versuchsperson der Median ermittelt und dann der Mann-Whitney-U-Test und der Kolmogorov-Smirnov-Test durchgeführt ($\alpha = 0,05$). Mediane der JFC-Werte des Termins A der Reihenfolge ANLT-JFCT und JFCT-ANLT zeigen keine signifikanten Unterschiede ($p_U = 0,699$, $p_{KS} = 0,81$). Das Ergebnis bleibt ähnlich, wenn auch die Werte des Termins B mitberücksichtigt werden ($p_U = 0,751$, $p_{KS} = 0,786$). Werden nur die Ergebnisse des Termins B untersucht so zeigt sich auch dabei kein signifikanter Unterschied ($p_U = 0,81$, $p_{KS} = 0,818$).

In Anhang D.4.9 sind die gesamten Datenwerte des ANLT und JFCT in den Reihenfolgen ANLT-JFCT und JFCT-ANLT als Boxplot dargestellt.

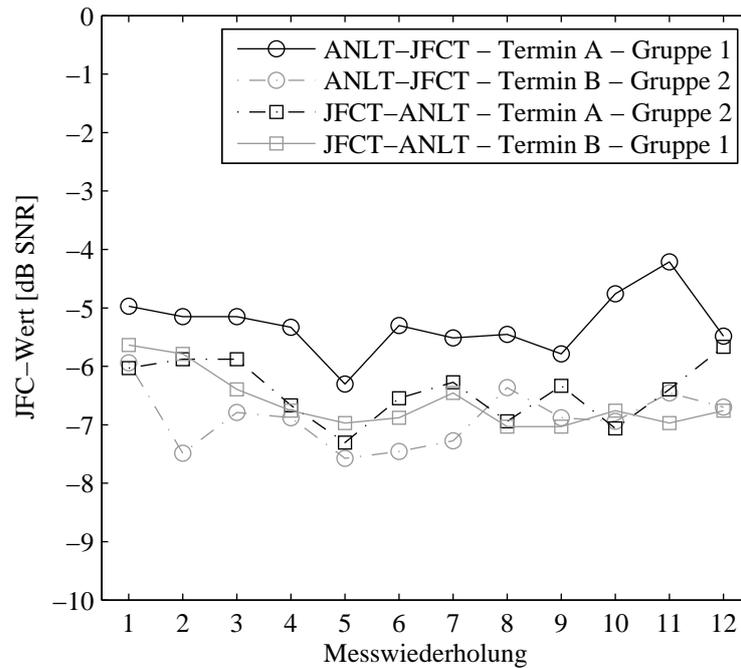


Abb. 3.11: Mediane der JFC-Werte. Dargestellt sind die Medianwerte, die für jede JFC-Messung mit den unterschiedlichen Parametern Reihenfolge, Termin und Versuchspersonengruppe ermittelt wurden.

Diskussion

Die Reliabilitätsmessungen des ANLT zeigten eine hohe Korrelation von 0,87 zwischen den Messwerten der Termine A und B. Diese Korrelation besitzt ein ähnliches Signifikanzniveau wie die von Nabelek *et al.* (2004) und Freyaldenhoven *et al.* (2006c) beschriebene. Eine Untersuchung der benötigten Messwiederholungen zur Bestimmung des ANL ergab kein eindeutiges Ergebnis. Es konnte jedoch beobachtet werden, dass mehrmalige Messwiederholung zu einem Anstieg des ANL führt, der vor allem durch eine Erhöhung des MCL hervorgerufen wird (s. Abb. 3.6). Einige Versuchspersonen beschrieben, dass die Einstellung ermüdend ist und sie einen höheren Sprachpegel im Laufe der Messung bevorzugten. Auf diesen Ergebnissen baut die Untersuchung in Kap. 3.8 auf. Sie untersucht, ob durch einen konstanten MCL konstantere ANL-Werte erreicht werden können.

Die Ergebnisse des JFCT zeigen eine geringere Korrelation von 0,69 zwischen den Ergebnissen der Termine A und B. Durch Latzel *et al.* (2003) wurde bereits eine geringe Korrelation beschrieben. Der JFCT eignet sich deshalb nicht für Wiederholungsmessungen an unterschiedlichen Tagen. Das JFC-Kriterium einer Versuchsperson verändert sich, wenn die Messungen mit zeitlichem Abstand durchgeführt werden. Wird die Messung direkt wiederholt, so kann keine Veränderung des JFC-Kriteriums festgestellt werden. Latzel *et al.* (2003) empfehlen deshalb, JFC-Werte innerhalb einer Sitzung zu messen. Um die Anzahl der notwendigen Messwiederholungen des JFCT zu ermitteln, wurden die JFC-Werte kumuliert und mit einem Friedman-Test geprüft, ob es signifikante Unterschiede zwischen den Kumulationen gibt. Ermittelt wurde, dass bei einer Kumulierung

von acht Messwiederholungen noch signifikante Unterschiede vorhanden sind. Daraus ist zu folgern, dass wenigstens neun Messwiederholungen durchgeführt werden sollten, da sie keine signifikanten Unterschiede mehr zeigen. Die Versuchspersonengruppe für dieses Experiment war jedoch klein und die Versuchspersonen jung. Bei Messungen mit älteren Versuchspersonen, für die die Handhabung einer Maus evtl. schwierig sein könnte, könnte es sinnvoll sein, zehn Messwiederholungen durchzuführen. Dies ist die maximale Anzahl, die in der Literatur gefunden werden kann (Neumann und Jensen, 2001). Der zusätzliche zeitliche Aufwand zur Messung eines weiteren JFC-Wertes ist gering und beträgt nicht mehr als eine Minute. Außerdem sollte die Handhabung der Maus im JFCT in einer weiteren Studie auch mit älteren Versuchspersonen untersucht werden.

Wird die Abhängigkeit der Messreihenfolge von JFCT und ANLT untersucht, können keine Unterschiede durch die verschiedenen Messreihenfolgen ermittelt werden. Die Darstellung der Messergebnisse zeigt aber sowohl beim ANL als auch beim JFC eine systematische Veränderung der Werte einer Testsituation (s. Abb. 3.10 und Abb. 3.11). Es ist jedoch nicht auszuschließen, dass es einen Einfluss der Messreihenfolge gibt, der weder durch die Versuchspersonengruppe, noch durch die Messreihenfolge oder den zeitlichen Ablauf erklärt werden kann.

Aus dem Experiment D sind folgende Schlussfolgerungen zu ziehen: Die Test-Retest-Reliabilität des ANLT von 0,87 ist hoch. Ihre Signifikanz ist mit den Ergebnissen von Nabelek *et al.* (2004) und Freyaldenhoven *et al.* (2006c) vergleichbar. Die Test-Retest-Reliabilität des JFCT von 0,69 ist geringer. Der JFCT eignet sich aufgrund dieser Tatsache nur für Messungen mit kurzem zeitlichem Abstand, und es sollten mindestens neun Wiederholungsmessungen durchgeführt werden.

3.8 Experiment E: Einfluss der Methodik bei Wiederholungsmessungen auf den ANL

In Kap. 3.7 wurde ein Anstieg des MCL bei mehreren Messwiederholungen des ANLT ermittelt, der einen Anstieg des ANL zur Folge hatte. Durchgeführt wurden diese Messungen indem eine MCL-Bestimmung von einer BNL-Bestimmung gefolgt wurde. Diese beiden Messungen wurden aufeinanderfolgend wiederholt. In der Literatur werden unterschiedliche Studien beschrieben, die unterschiedliche Vorgehensweisen bei mehrmaliger Ermittlung des ANL verwenden. Zum einen kann die MCL- und BNL-Bestimmung aufeinander folgend durchgeführt werden, d. h., dass den Versuchspersonen erst die Sprache und darauf folgend das Hintergrundgeräusch mit Sprache präsentiert wird (Nabelek *et al.*, 1991; Crowley und Nabelek, 1996; Harkrider und Tampas, 2006; Freyaldenhoven *et al.*, 2006e; Nabelek *et al.*, 2006; Freyaldenhoven *et al.*, 2006d; von Hapsburg und Bangh, 2006) (s. Kap. 3.3). Bei Messwiederholungen wird dieser gesamte Ablauf zusammenhängend wiederholt. Andere Studien verwenden bei Messwiederholungen eine andere Messreihenfolge (Rogers *et al.*, 2003; Freyaldenhoven *et al.*, 2005; Mueller *et al.*, 2006). Zuerst wird der MCL der Sprache einmal oder öfter bestimmt und gemittelt, um

danach die Bestimmung des BNL mehrmals durchzuführen. Der Sprachpegel wird dabei konstant beim zuvor gemittelten MCL präsentiert. Diese Vorgehensweise mit Trennung der MCL- und BNL-Bestimmung soll in diesem Experiment durchgeführt werden, und es soll geprüft werden, ob andere bzw. konstantere Ergebnisse erreicht werden als mit einer aufeinander folgenden MCL- und BNL-Bestimmung wie in Kap. 3.7.

Methoden

Der ANLT wurde an zwei verschiedenen Terminen (Termin A und B) mit einem Abstand von vier bis sieben Tagen durchgeführt. Jede Versuchsperson bestimmte zunächst dreimal den MCL. Diese Anzahl wurde gewählt in Übereinstimmung mit der Literatur (Rogers *et al.*, 2003; Harkrider und Tampas, 2006; Mueller *et al.*, 2006). Der Median der MCL-Werte wurde bestimmt und als Sprachpegel bei der Bestimmung des BNL-Wertes verwendet. Darauf folgend wurde die BNL-Bestimmung zwölfmal wiederholt.

Der Startpegel der Sprache bei der Bestimmung des MCL betrug immer 30 dB(A), der Startpegel des Hintergrundgeräusches bei der Ermittlung des BNL lag bei 20, 30 oder 40 dB(A), um einen Lerneffekt bei den durchgeführten Pegelinstellungen zu vermeiden. Die Reihenfolge der Startpegel bei der Bestimmung des BNL wurde randomisiert. Dieses Vorgehen wurde am Termin B wiederholt.

Versuchspersonen

An dieser Messung nahmen zwölf normalhörende Versuchspersonen (vier Frauen, acht Männer) im Alter zwischen 23 und 33 Jahren teil (mittleres Alter: 26 Jahre). Die Versuchspersonen hatten keine Erfahrung mit dem ANLT. Die Spanne des Hörvermögens ist in Anhang 3.8 dargestellt.

Ergebnisse

Die Ergebnisse wurden wie in Kap. 3.7 ausgewertet. Abb. 3.12 zeigt (a) die gemessenen MCL-Werte, (b) BNL-Werte und (c) ANL-Werte in Boxplots zusammenfasst. Auf der Abszisse sind die Wiederholungsmessungen und auf der Ordinate der Pegel in dB(A) und der ANL in dB SNR aufgetragen. Eine Erhöhung des ANL durch häufigere Wiederholung wie in Abb. 3.6c kann durch Abb. 3.12c nicht gezeigt werden. Deutlich wird jedoch, dass die Mediane des Termins B höhere ANL-Werte erreichen als die Werte des Termins A. Dieses Ergebnis ist signifikant nach einem Friedman-Test ($\chi^2_{(1;246;0,05)} = 15,672, p \approx 0$), in dem alle ANL-Werte der Termine A und B verglichen wurden.

Um den Zusammenhang zwischen den Ergebnissen des MCL, BNL und ANL beider Termine zu untersuchen, wurde für jede Versuchsperson der Median der zwölf Messwiederholungen zu den Terminen A und B ermittelt und der Spearmansche Korrelationskoeffizient r_{sp} bestimmt. Seine Irrtumswahrscheinlichkeit p wurde mit einem Permutationstest ermittelt. Die Korrelationskoeffizienten r_{sp} liegen für den MCL und den BNL

bei 0,94 und für den ANL bei 0,88. Sie sind signifikant auf einem Niveau von $\alpha = 0,01$ und erreichen damit höhere Werte als in Kap. 3.7. Eine Gegenüberstellung der MCL-, BNL- und ANL-Werte in einem Scatterplot zeigt Anhang D.5.1.

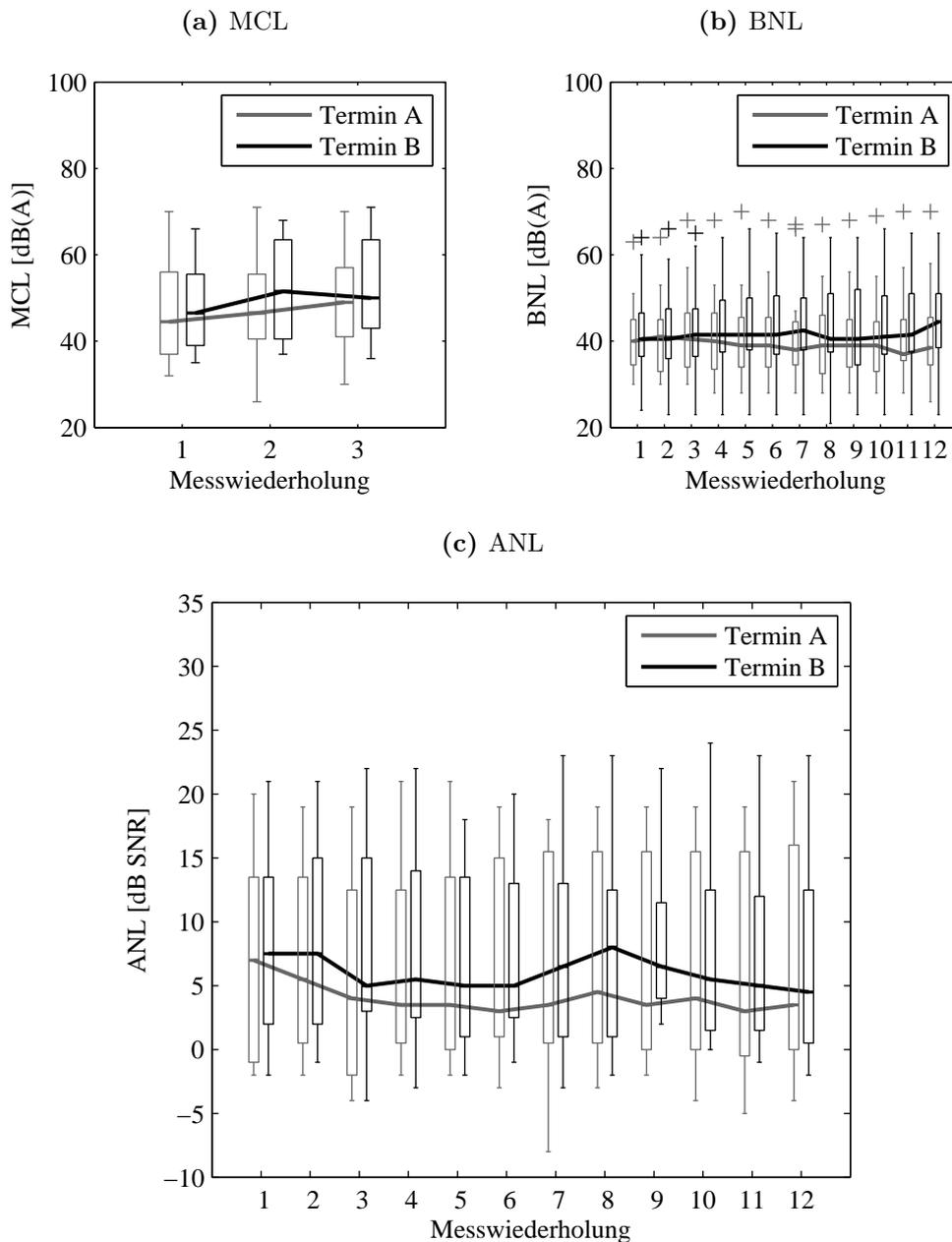


Abb. 3.12: (a) MCL-, (b) BNL- und (c) ANL- Werte, gemessen bei bis zu zwölf Messwiederholung in den Terminen A und B.

Die Streuung der ANL-Werte wurde mit Hilfe des Interquartilsabstandes bestimmt. Der Median beträgt 2,5 dB SNR (Termin A) bzw. 2 dB SNR (Termin B) und bleibt damit für den Termin A gleich und wird im Termin B geringer als in Kap. 3.7. Die maximale und minimale Streuung beträgt 6 dB SNR (Termin A) bzw. 4 dB SNR (Termin B) und 1 dB SNR (Termin A) bzw. 0 dB SNR (Termin B). Die Streuungswerte stehen im Gegensatz zu den Ergebnissen in Kap. 3.7, in denen im Termin B eine höhere Streuung erreicht wurde. Eine Darstellung der Streuungen findet sich in Anhang D.5.2.

Um die Anzahl der nötigen Wiederholungen der Messung zu ermitteln, wurde eine Kumulierung der Ergebnisse vorgenommen. Für die Abb. 3.13 wurde der Median der ersten ANL-Werte jeder Versuchsperson ermittelt. Kumuliert wurde die auf der Abszisse dokumentierte Anzahl der Werte. Ist die Anzahl der gemittelten Messungen z. B. drei, so wurde für jede Versuchsperson der Median der ANL-Werte, die sie bei den ersten drei Messungen eingestellt hatte, gebildet. Die Ergebnisse zeigen, dass die Werte im Verlauf der Messung nicht ansteigen wie in Abb. 3.7 (Kap. 3.7).

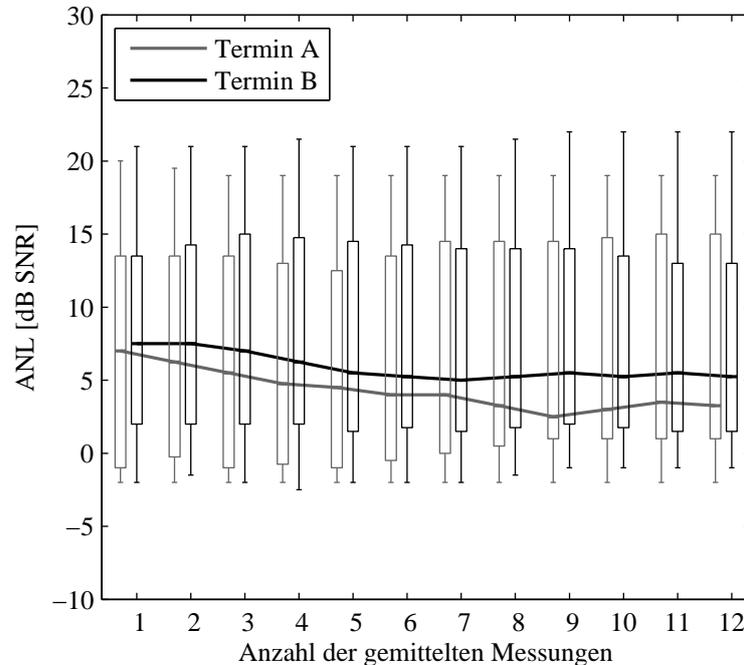


Abb. 3.13: Kumulierte ANL-Werte jeder Versuchsperson, gemessen in den Terminen A und B. Für jede Versuchsperson wurden aus der auf der Abszisse angegebenen Anzahl der ANL-Werte der Median ermittelt. Ist die Anzahl der gemittelten Messungen z. B. drei, so wurde für jede Versuchsperson der Median der Werte, die sie bei den ersten drei Messungen eingestellt hatte, gebildet.

Außerdem wurden die Unterschiede der kumulierten ANL-Werte mit Hilfe des Friedman-Tests auf einem Niveau von $\alpha = 0,05$ untersucht. Für den Vergleich zwischen den Medianen aus zwei und sechs ANL-Werten zeigt sich noch ein signifikanter Unterschied zwischen den Ergebnissen ($\chi^2_{(1;24;0,05)} = 4,75, p = 0,029$). Höhere Kumulationen zeigen keine Signifikanz mehr. Alle signifikanten Ergebnisse und eine von der interindividuellen Streuung befreite Betrachtung der Ergebnisse sind in den Anhängen D.5.3 und D.5.4 aufgeführt.

Um die ANL-Werte des Experimentes E mit denen des Experimentes D aus Kap. 3.7 zu vergleichen, wurden Mediane aus den zwölf Messwiederholungen für jede Versuchsperson ermittelt und diese Werte mit einem Kolmogorov-Smirnov-Test und einem Mann-Whitney-U-Test geprüft. Die Werte des Experimentes D und E zeigten keine signifikanten Unterschiede nach dem Kolmogorov-Smirnov-Test ($p = 0,987$) und ebenso nicht nach dem Mann-Whitney-U-Test ($p = 0,757$) bei einem Signifikanzniveau von $\alpha = 0,05$.

Diskussion

Die Ergebnisse des Experimentes E zeigen, dass durch eine Trennung der MCL-Messung von der BNL-Messung konstantere Ergebnisse erreicht werden können (s. Abb. 3.12) als durch eine mehrfache Wiederholung des gesamten Tests (s. Abb. 3.6). Ein Anstieg des ANL wie in Experiment D konnte nicht ermittelt werden. Sollen mehrere Messwiederholungen durchgeführt werden, ist es sinnvoller, zuerst den MCL zu ermitteln, die Sprache anschließend bei MCL zu präsentieren und dann folgend die BNL-Bestimmungen durchzuführen. Eine Aussage über die Anzahl der notwendigen Messwiederholungen kann anhand der kumulierten ANL-Werte getroffen werden. Die dazu untersuchten Daten sind signifikant unterschiedlich für einen Vergleich von zwei kumulierten ANL-Werten mit sechs kumulierten ANL-Werten. Außerdem zeigen die Ergebnisse der Abb. 3.13, dass der erste ANL sich von den folgenden unterscheidet. Deshalb sollten drei Messwiederholungen durchgeführt werden, da sich die Ergebnisse von nur zwei Wiederholungen noch signifikant von mehr Wiederholungen unterscheiden. Die Anzahl von drei Wiederholungen ist ein Kompromiss aus dem möglichen Messaufwand, der sich erhöht, wenn mehrere Wiederholungen durchgeführt werden und einer Erhöhung der Genauigkeit der Messung. Auch in der Literatur werden Studien beschrieben die maximal drei Messungen für jedes untersuchte Merkmal nutzen (Rogers *et al.*, 2003; Harkrider und Tampas, 2006; Mueller *et al.*, 2006; Franklin *et al.*, 2006).

Folgerung des Experimentes E ist, dass die Test-Retest-Reliabilität mit einer veränderten Methodik bei Wiederholungsmessungen erhöht werden kann. Dazu sollte bei Messwiederholungen zuerst der MCL öfter bestimmt und gemittelt werden, um danach die Bestimmung des BNL mehrmals durchzuführen. Eine Anzahl von drei Messwiederholungen im ANLT ist zu empfehlen.

3.9 Experiment F: Einfluss der SRT-Messung auf ANL- und JFC-Werte

Zur genaueren Prüfung der Zusammenhänge zwischen den Testverfahren wird in diesem Experiment der zeitlich unmittelbare Einfluss einer SRT-Messung, wie z. B. der des OLSA, auf den ANL und den JFC-Wert untersucht. Zu vermuten ist, dass ein Einfluss der SRT-Messung auf den JFC-Wert festgestellt werden kann, da OLSA und JFCT sich in ihrer Aufgabenstellung sehr ähnlich sind. Im OLSA wird eine Sprachverständlichkeitsschwelle objektiv ermittelt. Im JFCT hat die Versuchsperson die Aufgabe, eine subjektive Sprachverständlichkeitsschwelle einzustellen. Deshalb kann eine Versuchsperson bei einer Messreihenfolge, in der zuerst der OLSA und folgend der JFCT durchgeführt wird, den im OLSA eingestellten SNR als Anker verwenden und ihn im JFCT wiederholen. Die Aufgabenstellung des ANLT unterscheidet sich deutlich von Sprachverständlichkeitstests. Seine Ergebnisse sollten in einem geringeren Maß durch den OLSA beeinflusst werden als die des JFCT.

Eine Untersuchung des Einflusses einer SRT-Messung auf JFC- und ANL-Werte wur-

de in der Literatur bisher nicht beschrieben. In verschiedenen Studien wurde lediglich der Zusammenhang zwischen den Ergebnissen des ANLT oder des JFCT und einer SRT-Messung unter dem Einfluss unterschiedlicher Parameter untersucht. Verschiedene Studien untersuchten den Einfluss von z. B. Störgeräuschreduktionen im Hörgerät, monauralen oder binauralen Hörgeräteversorgungen und der Sprachkenntnisse eingewandelter koreanischer Versuchspersonen auf die Sprachverständlichkeit und den ANL (Nabelek *et al.*, 2004; Freyaldenhoven *et al.*, 2005; Nabelek *et al.*, 2006; von Hapsburg und Bangh, 2006; Freyaldenhoven *et al.*, 2006d,e; Mueller *et al.*, 2006). Geprüft wurde außerdem, ob mit dem ANLT eine Vorhersage der Sprachverständlichkeit, gemessen mit einem Sprachverständlichkeitstest, möglich ist (Nabelek *et al.*, 2004; Freyaldenhoven *et al.*, 2005; Nabelek *et al.*, 2006; von Hapsburg und Bangh, 2006; Mueller *et al.*, 2006). Die Studien von Nabelek *et al.* (2004, 2006), von Hapsburg und Bangh (2006) und Mueller *et al.* (2006) fanden keinen Zusammenhang zwischen den Messungen. Freyaldenhoven *et al.* (2005) dagegen zeigten einen Zusammenhang bei der Untersuchung des direktionalen Nutzens von Hörgeräten. Diese Ergebnisse weisen darauf hin, dass eine Sprachverständlichkeitsmessung vermutlich nur einen geringen Einfluss auf die Ergebnisse im ANLT nimmt. Dies könnte in der unterschiedlichen Aufgabenstellung der Tests begründet sein. Für den Zusammenhang zwischen JFC- und SRT-Werten zeigten sich in der Literatur unterschiedliche Ergebnisse. Larsby und Arlinger (1994) und Magnusson *et al.* (2001) ermittelten für unterschiedliche Störgeräusche und Versuchspersonen mit verschiedenen Graden an Schwerhörigkeit z. T. signifikante Zusammenhänge zwischen dem JFC-Wert und einem SRT 50%-iger Wortverständlichkeit. Latzel *et al.* (2003) und Neumann und Jensen (2001) beschrieben dagegen das Fehlen von Zusammenhängen zwischen JFC- und SRT-Werten. Zu vermuten ist, dass es einen Einfluss der SRT-Messung auf den JFC gibt, da die dabei präsentierten Instruktionen sehr ähnlich sind.

Methoden

Alle Versuchspersonen führten die Bestimmung des MCL, wie sie im ANLT üblich ist (s. Kap. 3.3), dreimal durch. Der Median des MCL wurde wie in Kap. 3.8 für die folgenden Messungen als Sprachpegel verwendet. Auch der BNL wurde anhand des ANLT-Verfahrens dreimal ermittelt. Der JFCT wurden zehnmal durchgeführt und der SRT 50%-iger Satzverständlichkeit wurde mit Hilfe einer Liste à 30 Sätzen des OLSA gemessen.

Um den Einfluss der SRT-Messung zu untersuchen, wurden die Messungen in einem Cross-Over-Design durchgeführt. Dazu wurden die Versuchspersonen in zwei Gruppen aufgeteilt. Gruppe 1 führte die Bestimmung des MCL, des JFC, des SRT und zum Schluss des BNL durch. Mit der Gruppe 2 wurden zuerst der MCL und danach der BNL, der SRT und abschließend der JFC ermittelt. Die Abfolge der Messungen ist in Abb. 3.14 verdeutlicht.

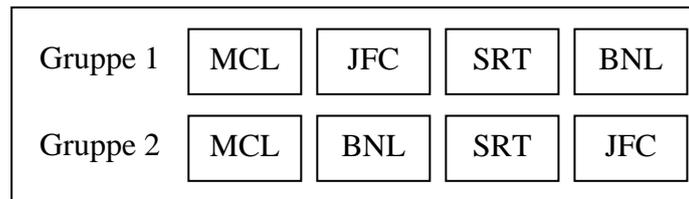


Abb. 3.14: Reihenfolge der unterschiedlichen Messungen für die Versuchspersonengruppen 1 und 2.

Versuchspersonen

An diesem Experiment nahmen zwölf Normalhörende (neun Frauen, drei Männer) teil. Ihre Hörschwellen sind als Spanne im Anhang C.6 dargestellt. Sie waren zwischen 21 und 25 Jahren alt (mittleres Alter: 23 Jahre). Alle Versuchspersonen hatten keine Erfahrung mit Hörtests. Sie wurden für ihren Aufwand mit 10 Euro/h entschädigt.

Ergebnisse

Zur Darstellung der Ergebnisse wurde der Median aus den drei Messwiederholungen des ANLT und der Median der zehn Messwerte des JFCT bestimmt und mit dem SRT des OLSA verglichen. Die Ergebnisse der beiden Versuchspersonengruppen wurden in Boxplots zusammengefasst und in Abb. 3.15 aufgetragen. Abb. 3.15a zeigt den JFC-Wert, Abb. 3.15b den SRT aus dem OLSA und 3.15c den ANL. Die Werte des JFCT und des OLSA sind sehr ähnlich. Sie liegen bei niedrigeren Pegeln und streuen weniger als die Werte des ANLT. Die ermittelten Mediane im JFCT und OLSA liegen zwischen -4,9 und -5,9 dB SNR in allen Gruppen. Im ANLT ist der Unterschied größer. Die Gruppe 1 (Reihenfolge: JFC, OLSA, BNL) erreicht einen ANL-Median von 4 dB SNR und die Gruppe 2 (Reihenfolge: BNL, OLSA, JFCT) einen Median von 7 dB SNR.

Um zu untersuchen, ob aufgrund der geänderten Messreihenfolgen die Versuchspersonengruppen zu unterschiedlichen Werten gelangen, wurden die Ergebnisse jeder Gruppe mit einem Mann-Whitney-U-Test hinsichtlich ihrer zentralen Tendenz und einem Kolmogorov-Smirnov-Test auf Unterschiede der gesamten Verteilung untersucht. Die ermittelten Irrtumswahrscheinlichkeiten p beider Signifikanztests sind in Tab. 3.1 dargestellt und wurden mit einem Schwellenwert von $\alpha = 0,05$ geprüft. Sie zeigen, dass die ANL- und SRT- Werte beider Versuchspersonengruppen nicht signifikant unterschiedlich hinsichtlich ihrer zentralen Tendenz sind und nicht aus unterschiedlichen Grundgesamtheiten stammen. Im JFCT erreicht die Gruppe 2 (Reihenfolge: BNL, OLSA, JFCT) niedrigere Werte als die erste (s. Abb. 3.15a). Die Signifikanztests zeigen für die JFC-Werte geringe Irrtumswahrscheinlichkeiten, die jedoch den Schwellenwert von $\alpha = 0,05$ noch nicht unterschreiten. Deshalb kann nicht angenommen werden, dass die JFC-Werte beider Gruppen unterschiedlichen Grundgesamtheiten entstammen.

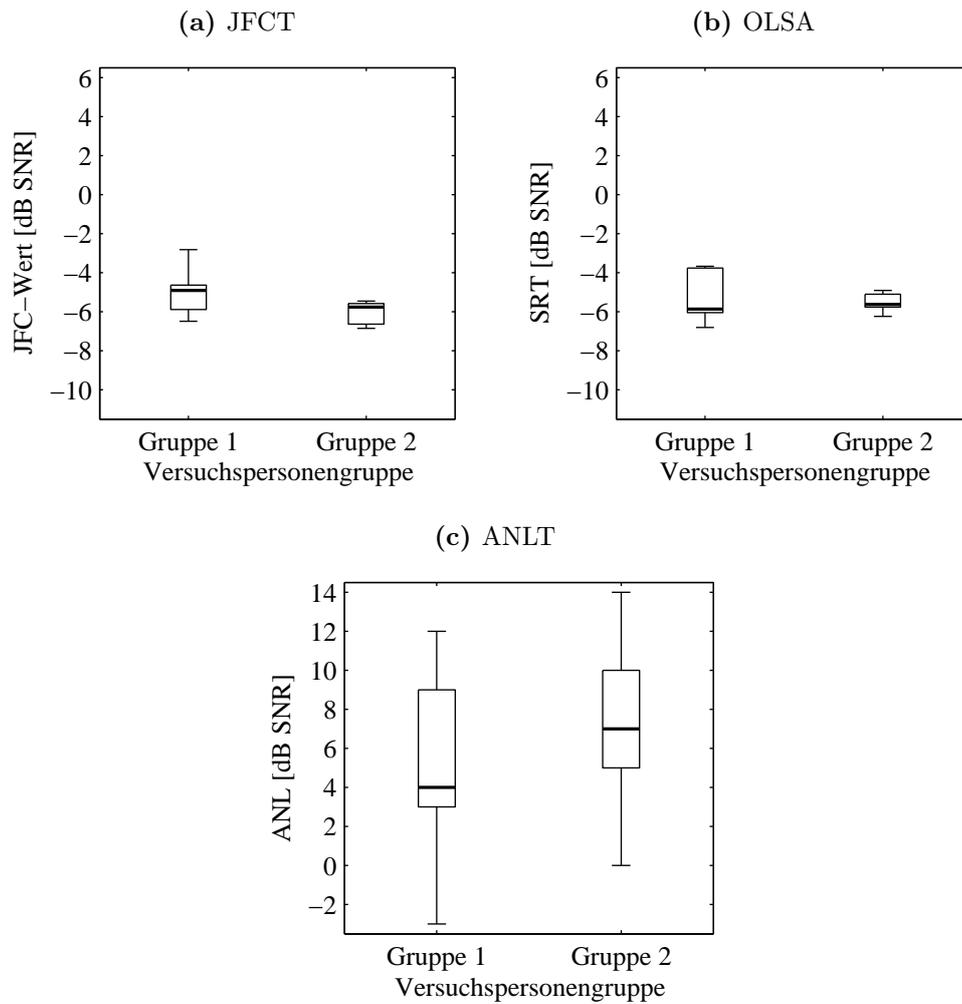


Abb. 3.15: Ergebnisse der Untersuchung des SRT-Einflusses auf den JFCT, ANLT. Abb. (a) zeigt den JFC-Wert, Abb. (b) zeigt den SRT gemessen im OLSA und Abb. (c) zeigt den ANL. Aufgetragen sind die Ergebnisse der Gruppe 1 (Reihenfolge: JFCT, OLSA, BNL) und der Gruppe 2 (Reihenfolge: BNL, OLSA, JFCT).

Tab. 3.1: Prüfung der signifikanten Unterschiede der Ergebnisse unterschiedlicher Versuchspersonengruppen im JFCT, OLSA und ANLT mit dem Mann-Whitney-U-Test und dem Kolmogorov-Smirnov-Test. Für jedes Untersuchungsverfahren ist die im Signifikanztest ermittelte Irrtumswahrscheinlichkeit p dargestellt.

| Untersuchungsverfahren | Mann-Whitney-U-Test | Kolmogorov-Smirnov-Test |
|------------------------|---------------------|-------------------------|
| JFC-Wert | $p = 0,087$ | $p = 0,077$ |
| SRT | $p = 0,31803$ | $p = 0,667$ |
| ANL | $p = 0,422$ | $p = 0,81$ |

Um zu prüfen, ob JFCT und OLSA zu unterschiedlichen Ergebnissen gelangen, wurden mit einem Wilcoxon Vorzeichen-Rangtest die SRT-Werte des OLSA und die JFC-Werte geprüft. Werden nur die Ergebnisse der ersten Gruppe (Reihenfolge: JFCT, OLSA, BNL) beider Testverfahren verglichen, so zeigt sich kein signifikanter Unterschied ($p = 0,5, \alpha = 0,05$). Die Unterschiede werden ebenfalls nicht signifikant, wenn die Ergebnisse beider Gruppen verglichen werden ($p = 0,519, \alpha = 0,05$).

Diskussion

In Experiment F wurden keine signifikanten Unterschiede der Ergebnisse der zwei Versuchspersonengruppen, die ANL-, JFC- und SRT-Messungen in unterschiedlichen Reihenfolgen durchführten (Gruppe 1: JFCT, OLSA, BNL; Gruppe 2: BNL, OLSA, JFCT), ermittelt (s. Tab. 3.1). Eine vorangestellte SRT-Ermittlung mit dem OLSA hat also keine Auswirkungen auf die Ergebnisse des ANLT. Für die JFC-Werte ist jedoch anhand der geringen Irrtumswahrscheinlichkeit im Signifikanztest der Trend zu erkennen, dass sie durch eine zuvor durchgeführte SRT-Messung beeinflusst werden können. Die Irrtumswahrscheinlichkeit, dass die Messwerte aus der gleichen Grundgesamtheit stammen, liegt bei 0,09, getestet mit einem Kolmogorov-Smirnov-Test. Sie sollte mit einer größeren Anzahl an Versuchspersonen überprüft werden. Außerdem zeigen die Ergebnisse des OLSA und die JFC-Werte keinen signifikanten Unterschied, d. h., dass die Versuchspersonen ähnlich gut die subjektive Sprachverständlichkeit des JFCT einregeln konnten wie die gemessene Sprachverständlichkeitsschwelle des OLSA. Für Messungen, in denen unterschiedliche Verfahren genutzt werden, ist zu empfehlen, den JFCT vor dem OLSA durchzuführen. So wird eine Beeinflussung ausgeschlossen.

Fazit des Experimentes F ist, dass eine SRT-Messung keinen Einfluss auf die ANL-Werte, aber tendenziell auf die JFC-Werte hat. Deshalb sollte der JFCT vor einer SRT-Messung durchgeführt werden.

3.10 Experiment G: Einfluss unterschiedlicher Hintergrundgeräusche

Mehrere Studien zeigten eine Abhängigkeit des ANL vom verwendeten Hintergrundgeräusch (Nabelek *et al.*, 1991; Freyaldenhoven *et al.*, 2006c; Gordon-Hickey und Moore, 2007). Nabelek *et al.* (1991) verwendeten a) Stimmengewirr von zwölf Sprechern aus dem Speech Perception in Noise Test (Kalikov *et al.*, 1977), b) Rauschen mit der gleichen spektralen Einhüllenden der verwendeten Sprache, c) eine Aufnahme von Verkehrsgereuschen, d) leise Musik und e) die Aufnahme eines Pressluftbohrers. Die Ergebnisse zeigten, dass Musik im Vergleich zu anderen Hintergrundgeräuschen zu unterschiedlichen Ergebnissen führen kann. Dabei ist auch die Art der Musik zu berücksichtigen. Diese Ergebnisse wurden von Gordon-Hickey und Moore (2007) aufgegriffen. Sie untersuchten den Einfluss verschiedener Musik und des Musikgeschmacks auf den ANL bei Normalhörenden. Die Ergebnisse zeigten, dass der ANL nicht mit dem Musikgeschmack

zusammenhängt und ANL-Werte, gemessen mit Stimmengewirr, höher liegen als mit Musik. Die Versuchspersonen tolerierten also einen geringen Pegel bei Stimmengewirr als bei Musik. Freyaldenhoven *et al.* (2006c) untersuchten nochmals den Unterschied zwischen Stimmengewirr von zwölf Personen und einem sprachsimulierendem Rauschen. Der mit sprachsimulierendem Rauschen ermittelte ANL-Wert lag bei Normalhörenden signifikant höher als mit Stimmengewirr. D. h., die Versuchspersonen tolerierten einen geringeren Pegel des sprachsimulierenden Rauschens als des Stimmengewirrs. Dieses Experiment greift die Zielsetzungen der zuvor genannten Studien auf und untersucht den ANL mit unterschiedlichen Hintergrundgeräuschen.

Methoden

Untersucht wurden fünf verschiedene Hintergrundgeräusche: a) sprachsimulierendes Rauschen (OLSA-Rauschen), b) die Aufnahme einer Maschine (Maschine), c) eine Aufnahme in einer Cafeteria (Cafeteria), d) Stimmengewirr von zwölf unterschiedlichen englischsprachigen Sprechern unterschiedlichen Geschlechts (ANL-Stimmengewirr, Kalikov *et al.* (1977)) und e) ein Stimmengewirr von weiblichen Sprechern (weibliches Stimmengewirr). Das Hintergrundgeräusch e) wurde aus dem International Speech Test Signal (ISTS, Holube *et al.* (2007)) entwickelt. Dieses Signal besteht aus Aufnahmeausschnitten von sechs Sprecherinnen. Die Sprecherinnen waren unterschiedlicher Nationalität und sprachen ihre Muttersprache (Arabisch, Chinesisch, Deutsch, Englisch, Französisch und Spanisch). Das ISTS wurde sechsmal addiert, um ein Stimmengewirr zu erhalten. Für jede Addition wurde ein Signalausschnitt gewählt, dessen Beginn um einen Betrag verschoben war, der 20 s plus einem randomisierten Wert entsprach. Durch dieses Vorgehen entstanden Signalausschnitte, die nicht immer die benötigte Länge von 1 min besaßen. An zu kurze Signalausschnitte wurden Signalteile vom Beginn des ISTS bis zur erforderlichen Länge angehängt. So entstand ein einminütiges Stimmengewirr. Die Spektren der Rauschen sind in Abb. 3.16 dargestellt.

Die Versuchspersonen bestimmten in dieser Untersuchung zuerst dreimal den MCL. Aus diesen Werten wurde ein Median ermittelt, der in der folgenden Messung als konstanter Sprachpegel für die Sprache verwendet wurde (s. dazu Kap. 3.8). Anschließend wurde mit jedem Hintergrundgeräusch der ANL dreimal ermittelt. Es wurde also eine Gesamtanzahl von 15 ANL-Bestimmungen durchgeführt. Die Reihenfolge der Hintergrundgeräusche wurde für jede Versuchsperson randomisiert.

Versuchspersonen

Diese Messung führten zwölf normalhörende Versuchspersonen (fünf Frauen, sieben Männer) im Alter zwischen 24 und 34 Jahren (mittleres Alter: 28,2 Jahre) durch. Ihr Hörvermögen ist in C.7 dargestellt. Alle Versuchspersonen hatten Erfahrung mit dem ANLT, denn sie hatten an mindestens einem der Experimente A-D teilgenommen.

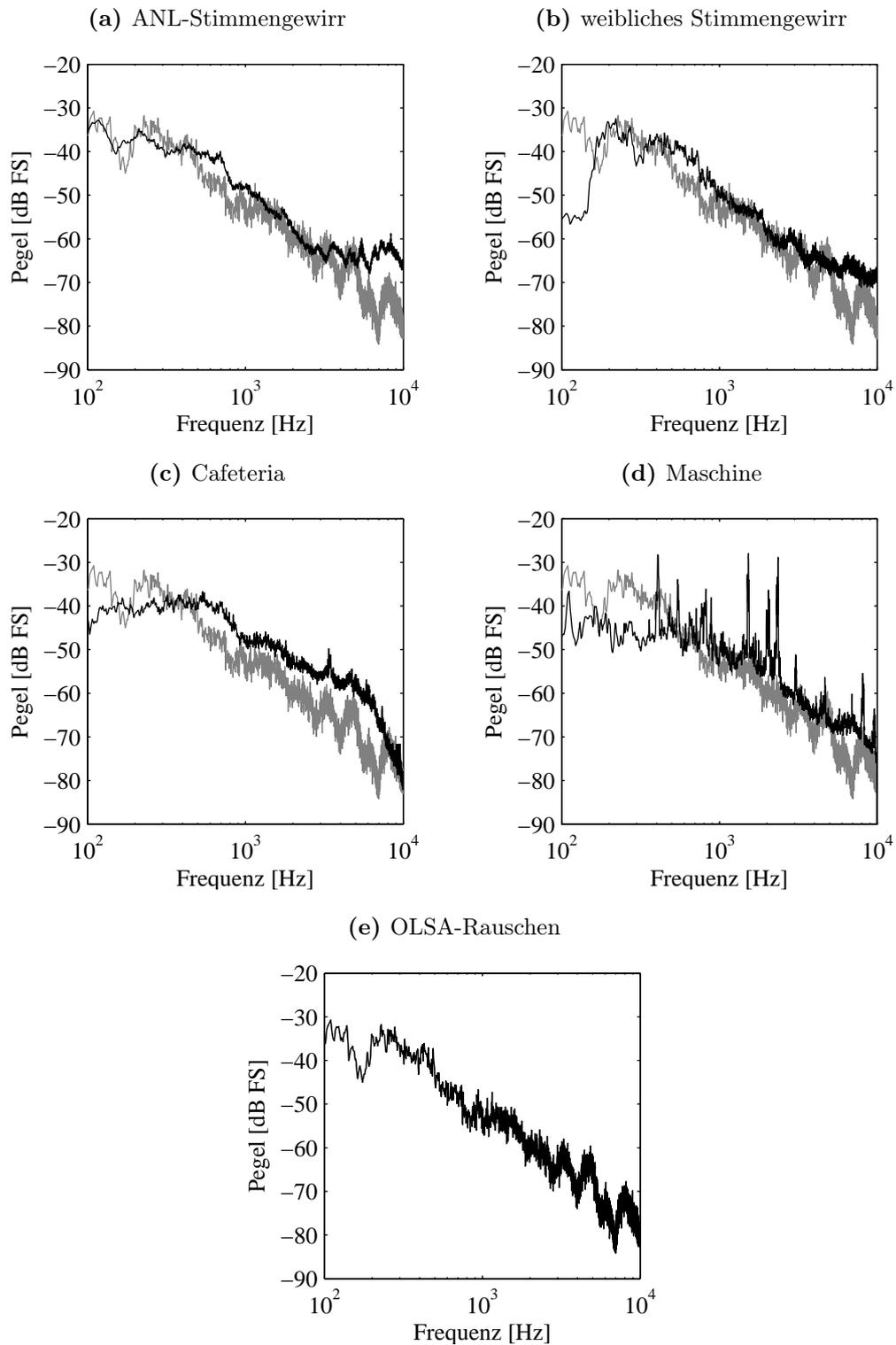


Abb. 3.16: Spektren der verwendeten Rauschen. (a) zeigt das Spektrum des Stimmengewirrs von zwölf Sprechern, (b) des weiblichen Stimmengewirrs, (c) der Aufnahme in der Cafeteria, (d) der Maschine und (e) des OLSA-Rauschens. In hellgrau ist jeweils das Spektrum der verwendeten Sprache hinterlegt.

Ergebnisse

Zur Darstellung der ANL-Werte wurde der Median der drei Messwiederholungen für jede Versuchsperson ermittelt und in Abb. 3.17 dargestellt. Die beiden Stimmengewirre (ANL-Stimmengewirr, weibliches Stimmengewirr) erreichen die höchsten medianen ANL-Werte mit 13,5 dB SNR und 11,5 dB SNR. Ein Friedman-Test ergab, dass es signifikante Unterschiede zwischen den ANL-Werten mit unterschiedlichen Hintergrundgeräuschen gibt ($\chi^2_{(4;44;0,05)} = 19,155$, $p \approx 0$).

Zur Untersuchung der Unterschiede zwischen den einzelnen Hintergrundgeräuschen wurde der Wilcoxon Vorzeichen-Rang-Test angewendet. Anhang D.6 zeigt die ermittelten Irrtumswahrscheinlichkeiten. Das ANL-Stimmengewirr ist signifikant unterschiedlich zu den anderen Hintergrundgeräuschen bei einem Signifikanzniveau von $\alpha = 0,05$. Wird das Signifikanzniveau nach Bonferroni (Bortz *et al.*, 2000) adjustiert, beträgt es $\alpha = \frac{0,05}{5} = 0,01$. Danach sind nur noch die Werte gemessen mit ANL-Stimmengewirr und OLSA-Rauschen und ANL-Stimmengewirr und Cafeteria signifikant unterschiedlich.

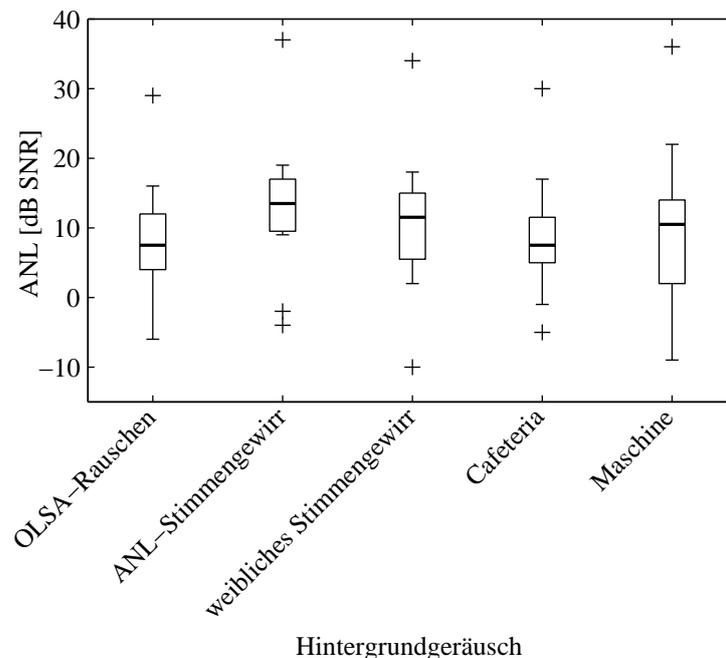


Abb. 3.17: ANL-Werte, gemessen mit unterschiedlichen Hintergrundgeräuschen.

Diskussion

Die Ergebnisse zeigen, dass der ANL abhängig ist von bestimmten Hintergrundgeräuschen (s. Abb. 3.17). Ein Unterschied der ANL-Werte wurde für ein Stimmengewirr von zwölf Sprechern beiderlei Geschlechts ermittelt. In diesem Fall stellten die Versuchspersonen signifikant höhere ANL-Werte ein. Sie tolerierten also weniger Hintergrundgeräusch. Stimmengewirr lenkt wahrscheinlich in einem größeren Maße von der Sprache ab als gleichförmige Geräusche, wie z. B. eine Maschine. Das ANL-Stimmengewirr ist für den Zuhörer durch seine zwölf Sprecher beiderlei Geschlechts dem Sprachmaterial

anscheinend am ähnlichsten und wird deshalb am wenigsten akzeptiert. Das weibliche Stimmengewirr besteht nur aus sechs weiblichen Sprechern und lenkt die Versuchspersonen vermutlich von dem männlichen Sprecher des OLSA-Materials nicht in dem Maße ab, wie das ANL-Stimmengewirr. Folgende Annahme könnte für die hier beschriebenen Ergebnisse gelten: Je unähnlicher die Hintergrundgeräusche dem Sprachsignal sind, desto eher könnten sie akzeptiert werden und desto niedriger sind die gewählten ANL-Werte.

In diesem Experiment wurde ermittelt, dass mit dem ANL-Stimmengewirr höhere ANL-Werte eingestellt werden als mit einem sprachsimulierenden Rauschen. Dieses Ergebnis wird von Nabelek *et al.* (1991) und Freyaldenhoven *et al.* (2006c) nicht bestätigt. Dort wird ein niedrigerer ANL mit dem Stimmengewirr erreicht als mit einem sprachsimulierenden Rauschen. Die absoluten Differenzen des mittleren ANL gemessen mit Stimmengewirr und Rauschen betragen 2,1 dB SNR (Freyaldenhoven *et al.*, 2006c) und 0,68 dB SNR (Nabelek *et al.*, 1991). Die hier ermittelte absolute Differenz zwischen den Medianwerten des ANL beträgt 6 dB SNR.

Ergebnis des Experimentes G ist, dass die Akzeptanz des Stimmengewirrs von zwölf Sprechern durch die Versuchspersonen geringer ist, als die Akzeptanz der anderen untersuchten Hintergrundgeräusche. Dieses Ergebnis stimmt nicht mit den Ergebnissen von Nabelek *et al.* (1991) und Freyaldenhoven *et al.* (2006c) überein.

3.11 Experiment H: Einfluss des Darbietungspegels bei der Übung des Sprachmaterials

Für den OLSA wird die Durchführung einer Trainingsphase empfohlen (Wagener *et al.*, 1999a,b,c; Wagener, 2003). Diese verhindert das Auftreten eines Lerneffektes während der Messung. Um einen Lerneffekt im ANLT zu vermeiden, wurde auch in den Experimenten A-G eine Trainingsphase durchgeführt. Die Versuchspersonen hörten maximal drei Listen des OLSA bei einem Pegel von 40 dB(A) und hatten die Aufgabe, die verstandenen Wörter zu wiederholen. Diese Übung des Sprachmaterials bei einem festen Pegel könnte Einfluss auf die Bestimmung des MCL im ANLT genommen haben. Versuchspersonen könnten den Pegel von 40 dB(A) als Anker verwendet haben und ihn im ANLT als MCL einstellen. Nach Franklin *et al.* (2006), Freyaldenhoven *et al.* (2006a) und Freyaldenhoven *et al.* (2006b) hängt der ANL signifikant vom dem Darbietungspegel der Sprache ab. Bei höherem Sprachpegel werden höhere ANL-Werte erreicht. Die Orientierung, die durch die Trainingsphase bei festem Darbietungspegel für die Sprache entstehen könnte, sollte jedoch im regulären ANLT durch die Einstellung eines minimalen und maximalen Sprachpegels aufgehoben werden. Dieses Experiment prüft den Einfluss des Darbietungspegels der Übungssätze auf den MCL und den ANL.

Methoden

Durchgeführt wurde eine Trainingsphase, in der zwei Listen mit 30 Sätzen des OLSA vorgespielt wurden. Der Darbietungspegel jedes einzelnen Satzes wurde randomisiert. Gewählt wurden für Normalhörende Darbietungspegel zwischen 35 und 65 dB(A) mit 5 dB-Differenz. Die präsentierten Listen wurden zufällig gewählt, und die Versuchsperson hatte die Aufgabe, die verstandenen Wörter zu wiederholen. Anschließend wurde der ANL gemessen. Dazu wurde die Ermittlung des MCL mit dem Sprachsignal dreimal durchgeführt, der Median des MCL berechnet und danach die Sprache zusammen mit dem Hintergrundgeräusch präsentiert. Auch hier wurde die Ermittlung des BNL dreimal vorgenommen.

Versuchspersonen

Versuchspersonen waren zwölf Normalhörende (6 Frauen, 6 Männer) im Alter zwischen 20 und 35 Jahren (mittleres Alter: 25,7 Jahre). Ihr Hörvermögen gemessen mit einem Tonaudiogramm ist Anhang C.8 dargestellt. Alle Versuchspersonen hatten keine Erfahrung mit dem ANLT und mit anderen Hörtests. Sie wurden für ihren Aufwand mit 10 Euro/h entschädigt.

Ergebnisse

Aus den drei Messwiederholungen aller Messwerte im ANLT wurde für jede Versuchsperson der Median bestimmt. Die Ergebnisse aller Versuchspersonen wurden dann in Boxplots der Abb. 3.18 zusammengefasst. Sie sind mit „randomisierter Pegel“ gekennzeichnet, da die Trainingsphase mit randomisierten Pegeln durchgeführt wurde. Außerdem wurden die Ergebnisse des Termins A der Untersuchung aus Kap. 3.8 zum Vergleich herangezogen. Aus den ersten drei Messwerten jeder Versuchsperson wurde ebenfalls der Median bestimmt und zusätzlich als Boxplots in der Abb. 3.18 dargestellt. Die Ergebnisse des Kap. 3.8 sind mit „konstanter Pegel“ gekennzeichnet, da in der Trainingsphase die Sätze mit einem konstanten Pegel von 40 dB(A) dargeboten wurden. Die Abb. 3.18 zeigt, dass die Mediane und die Streuungen aller Messungen im ANLT ähnlich sind. Die MCL-Werte liegen bei einem Median von 51 dB(A) bei einer Trainingsphase mit konstantem Pegel und bei 46,5 dB(A) bei einer Trainingsphase mit randomisiertem Pegel. Die Mediane der ANL-Werte bei randomisiertem Pegel in der Trainingsphase betragen 7 dB SNR und bei konstantem Pegel in der Trainingsphase 5,5 dB SNR. Die Streuung der ANL-Werte könnte auf mögliche signifikante Unterschiede hinweisen.

Die Messwerte beider Untersuchungen wurden mit einem Mann-Whitney-U-Test hinsichtlich ihrer zentralen Tendenz und mit einem Kolmogorov-Smirnov-Test hinsichtlich ihrer gesamten Verteilung geprüft. Die ermittelten Irrtumswahrscheinlichkeiten der Messwerte des ANLT sind in Anhang D.7 dargestellt. Beide Tests zeigen keine signifikanten Unterschiede des MCL und des ANL bei einem Signifikanzniveau von $\alpha = 0,05$.

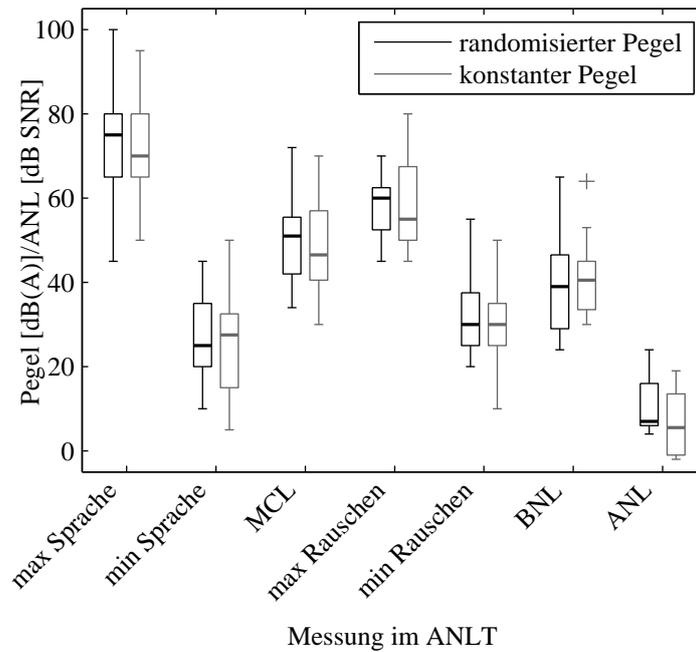


Abb. 3.18: Messwerte im ANLT bei der Untersuchung des Einflusses der Übung. In schwarz sind die Ergebnisse dieses Experiments mit der Trainingsphase bei randomisiertem Pegel und in grau die Ergebnisse aus Kap. 3.8 mit der Trainingsphase bei konstantem Pegel dargestellt.

Diskussion

Die Ergebnisse zeigen keine signifikanten Unterschiede aller Messwerte im ANLT, wenn die Übung des Sprachmaterials mit konstantem Pegel oder randomisiertem Pegel durchgeführt wird. Eine Übung des Sprachmaterials kann also bei einem konstanten Pegel erfolgen. Für Messungen mit Schwerhörigen ist dies von Vorteil, da aufgrund von Rekrutment Schwierigkeiten entstehen, unterschiedliche Pegel zur Darbietung der Sätze zu ermitteln.

Ergebnis des Experimentes H ist, dass eine Übung des verwendeten Sprachmaterials bei einem leisen konstanten Pegel vor Beginn des ANLT keinen Einfluss auf den ANL hat.

3.12 Schlussfolgerungen

Die Schlussfolgerungen der Experimente A-H können wie folgt zusammengefasst werden:

1. Mit der Schrittgröße 5 bzw. 2 dB kann der ANL schneller und präziser ermittelt werden als mit den Schrittgrößen 5 bzw. 1 dB und 2 bzw. 1 dB.
2. Die Verwendung einer auf- oder absteigenden Instruktion im ANLT führt zu unterschiedlichen Messergebnissen des MCL und BNL. Der ANL, die Differenz des MCL und BNL, wird jedoch nicht beeinflusst.
3. Geringe und mittlere Startpegel von 30 bzw. 50 dB(A) haben keinen Einfluss auf den ANL. Hohe Startpegel von 70 dB(A) weisen die Tendenz auf, den ANL zu hö-

heren Pegeln zu verschieben. Bei Messwiederholungen ist damit eine Veränderung des Startpegels möglich, um routiniertem Antwortverhalten vorzubeugen.

4. Die Test-Retest-Reliabilität des ANL ist mit $r_{sp} = 0,87$ recht hoch und ihre Signifikanz ist vergleichbar mit Ergebnissen von Nabelek *et al.* (2004) und Freyaldenhoven *et al.* (2006c).
5. JFC-Messungen eignen sich nur für Messungen mit kurzem zeitlichen Abstand und bei mindestens neun Messwiederholungen.
6. Die Test-Retest-Reliabilität kann erhöht werden, indem bei Messwiederholungen zuerst der MCL bestimmt wird, dann die Ergebnisse des MCL gemittelt werden und darauf die BNL-Bestimmungen bei gleichem mittleren MCL durchgeführt werden.
7. Drei Messwiederholungen sind zur Ermittlung des ANL zu empfehlen.
8. Eine SRT-Messung hat keinen Einfluss auf ANL-Werte, aber tendenziell auf JFC-Werte. Deshalb sollte der JFCT vor einer SRT-Messung durchgeführt werden.
9. ANL-Werte, gemessen mit einem Stimmengewirr von zwölf Sprechern, liegen höher als mit anderen Hintergrundgeräuschen. Diese Ergebnisse sind nicht vergleichbar mit den Ergebnissen von Nabelek *et al.* (1991) und Freyaldenhoven *et al.* (2006c). Diese Studien zeigen niedrigere ANL-Werte.
10. Eine Übung des verwendeten Sprachmaterials bei einem konstanten leisen Darbietungspegel vor Beginn des ANLT hat keinen Einfluss auf den ANL.

Die dargestellten Ergebnisse zeigen, dass der ANLT auf eine ähnliche Weise wie von Nabelek *et al.* (2006) vorgeschlagen, durchgeführt werden kann. Bei Wiederholungsmessungen sollte zuerst der MCL und dann folgend er BNL bestimmt werden. Der ANL ist damit ein zuverlässiges Maß zur subjektiven Beurteilung von Sprache im Störgeräusch.

In weiteren Messungen sollten alle Fragestellungen nochmals mit einer größeren Anzahl an Versuchspersonen geprüft werden. Da der ANL große interindividuelle Unterschiede zeigt, könnten mehr Versuchspersonen zu aussagekräftigeren Ergebnissen führen. Außerdem sollten z.B. die Fragestellungen der Schrittgröße und des Hintergrundgeräusches auch mit Ergebnissen von Schwerhörigen verglichen werden. Zusätzlich sollte untersucht werden, wie sich der ANL bei der Darbietung über Kopfhörer und Lautsprecher verhält.

Kapitel 4

Untersuchung von einkanaligen Störgeräuschreduktionen mit dem Acceptable Noise Level Test und anderen perzeptiven Maßen

4.1 Zusammenfassung

Diese Studie prüft die Eignung des Acceptable Noise Level Tests (ANLT) zur Untersuchung von einkanaligen Störgeräuschreduktionen. Betrachtet wurden drei einkanalige Störgeräuschreduktionen. Ein Algorithmus verwendete ein a priori Wissen des Hintergrundgeräusches und eine Verstärkungsfunktion nach Wiener und seine maximale theoretische Reduktion wurde auf 6 dB SNR beschränkt. Zwei weitere Algorithmen schätzten das Rauschen mit Hilfe des Minima Controlled Recursive Averaging und filterten das Signal nach der Methode der Spektralen Subtraktion der Leistungsdichtespektren. Ihre maximale theoretische Reduktion wurde auf 6 bzw. 8 dB SNR begrenzt. Die Ergebnisse zeigen für zehn normalhörende und elf schwerhörige Versuchspersonen, dass der ANLT zur Untersuchung von Störgeräuschreduktionen verwendet werden kann. Problematisch ist die interindividuelle Streuung des Acceptable Noise Level (ANL). Zum Vergleich wurden der ANLT, der Oldenburger Satztest (OLSA), der Just Follow Conversation Test (JFCT) und ein Paarvergleich durchgeführt. OLSA und JFCT sind zur Dokumentation des Nutzens von Störgeräuschreduktionen nicht geeignet, da im Test SNR-Werte erreicht werden, bei denen die Störgeräuschreduktion keine Verbesserung erzielt. Der Paarvergleich zeigt, dass die Schwerhörigen deutlich Signale, verarbeitet mit Störgeräuschreduktion, bevorzugen. Diese Aussage trifft nicht auf die Ergebnisse der Normalhörenden zu, die Störgeräuschreduktionen mit Artefakten meiden. Ein Vergleich der Ergebnisse des ANLT mit dem Paarvergleich zeigt, dass die Akzeptanz des Hintergrundgeräusches bei den Schwerhörigen z. T. von ihrer Bevorzugung der Störgeräuschreduktionen vor Signalen ohne Störgeräuschreduktion abhängig ist. Ihre Ergebnisse zeigen auch signifikante Korrelationen von 0,44-0,74 zwischen den ermittelten Ergebnissen der Sprachverständ-

lichkeitstests und des ANLT. Der akzeptable Rauschpegel ist für diese Versuchspersonengruppe u. a. abhängig von der Sprachverständlichkeit.

4.2 Einleitung

Ziel einkanaliger Störgeräuschreduktionen ist eine Verbesserung der Sprachverständlichkeit und/oder der Klangqualität durch Absenkung des Störsignals. Daher werden sie oft bei der Verarbeitung von Sprache im Hintergrundgeräusch eingesetzt und z. B. in Hörgeräten verwendet (Bentler und Chiou, 2006). Einkanalige Störgeräuschreduktionen bestehen in der Regel aus einer Störgeräuschschätzung und einer Verstärkungsregel. Die Störgeräuschschätzung stellt im gestörten Sprachsignal den Hintergrundgeräuschanteil fest. Eine Möglichkeit der Schätzung ist das „Minima Controlled Recursive Averaging“ von Cohen und Berdugo (2002). Es schätzt anhand der Minima des Signals in Sprachpausen das Rauschen. In einem zweiten Schritt wird die frequenzabhängige Verstärkung berechnet. Berechnungsformeln dafür sind z. B. das Wiener Filter oder die Spektrale Subtraktion (Vary *et al.*, 1998). Diese Verstärkungsregeln reduzieren den Pegel in verschiedenen Frequenzbändern, die vom Hintergrundgeräusch dominiert werden.

Eine naheliegende perzeptive Methode zur Untersuchung der Wirkung von Störgeräuschreduktionen ist die Bestimmung der Sprachverständlichkeit im Störgeräusch. Ziel dieses Tests ist die Ermittlung eines Signal-Rausch-Verhältnisses (engl. *Signal-to-Noise Ratio*, *SNR*), bei dem 50% der Sprache verstanden wird. Das erreichte Verhältnis wird als Sprachverständlichkeitsschwelle (engl. *Speech Reception Threshold*, *SRT*) bezeichnet. Die Bestimmung des SRT kann mit Word- oder Sentence-Scoring durchgeführt werden. Im Word-Scoring werden die verstandenen Wörter jedes Satzes gezählt und danach einen SNR-Wert für 50%-iges Wortverstehen eingeregelt. Beim Sentence-Scoring gilt ein Satz bereits als falsch, wenn nur ein Wort falsch wiederholt wird (Plomp und Mimpen, 1979; Nilsson *et al.*, 1994; Fredelake, 2006). Ermittelt wird ein SNR-Wert für 50%-iges Satzverstehen. Dieses Vorgehen führt zu einer Erhöhung der Schwierigkeit für die Versuchsperson, die höhere SNR-Werte als beim Word-Scoring benötigt, um 50%-iges Satzverstehen zu erreichen. Verschiedene Studien, in denen der SRT bestimmt wurde, zeigten bei der Untersuchung unterschiedlicher einkanaliger Störgeräuschreduktionen keine (Eberling *et al.*, 1993; Boymans *et al.*, 1999; Marzinzik, 2000; Walden *et al.*, 2000; Alcántara *et al.*, 2003; Ricketts und Hornsby, 2005; Mueller *et al.*, 2006) oder nur eine geringe Verbesserung der Sprachverständlichkeit (Bray und Nilsson, 2001; Lurquin *et al.*, 2001; Bray *et al.*, 2002). Erklärt wurden diese Ergebnisse u. a. mit der Untersuchungsmethode, dem Sprachverständlichkeitstest. SRT-Werte liegen vor allem für Normalhörende bei negativen SNR-Werten (Kollmeier und Wesselkamp, 1997; Wagener *et al.*, 1999a,b,c). Dabei hat das Hintergrundgeräusch einen höheren Pegel als die Sprache. Einkanalige Störgeräuschreduktionen arbeiten jedoch am effektivsten bei positiven SNR-Werten (Marzinzik, 2000).

Zur Untersuchung der Verbesserung der Klangqualität bei Verwendung einkanaliger Störgeräuschreduktionen werden häufig Paarvergleiche durchgeführt (Boymans und Dreschler, 2000; Ricketts und Hornsby, 2005). Dabei werden Signalaare dargeboten

und die Versuchspersonen haben z. B. das Signal mit der besseren Klangqualität zu wählen. Dieses Verfahren ist auch für geringe wahrnehmbare Unterschiede sensitiv. Da beim Paarvergleich jedes Signal mit jedem anderen Signal verglichen werden muss, um eine Rangordnung der Signale zu erhalten, steigt die Anzahl der Vergleiche mit der Anzahl der verwendeten Signale schnell an.

Neben der Klangqualität und der Sprachverständlichkeit nehmen einkanalige Störgeräuschreduktionen vermutlich Einfluss auf die Akzeptanz von Hintergrundgeräuschen. Zur Untersuchung kann der Acceptable Noise Level Test (ANLT) (Nabelek *et al.*, 1991, 2006) verwendet werden. Dieses Verfahren stellt den akzeptierbaren SNR für Sprache im Hintergrundgeräusch fest. Der so ermittelte SNR wird Acceptable Noise Level (ANL) genannt. Er ist definiert als Differenz zwischen dem Most Comfortable Level (MCL) der Sprache und dem höchsten Pegel des Hintergrundgeräusches, den die Versuchsperson über längere Zeit beim Hören von Sprache akzeptieren würde. Dieser Pegel des Hintergrundgeräusches wird Background Noise Level (BNL) genannt. Versuchspersonen mit einem niedrigen ANL tolerieren höhere Hintergrundgeräuschpegel als Versuchspersonen mit einem hohen ANL. Verschiedene Studien zeigten, dass der ANL unabhängig vom Grad des Hörverlustes, vom Alter, vom Geschlecht der Versuchsperson und von deren Nationalität (englisch, koreanisch) ist und auch davon, ob die Messungen mit oder ohne Hörgerät durchgeführt wurden (Nabelek *et al.*, 1991; Rogers *et al.*, 2003; Nabelek *et al.*, 2004; von Hapsburg und Bangh, 2006; Nabelek *et al.*, 2006). Der ANL war ebenso für die meisten Versuchspersonen nicht abhängig davon, ob monaural oder binaural mit Hörgeräten versorgt wurde (Freyaldenhoven *et al.*, 2006d). Außerdem zeigten Nabelek *et al.* (2004, 2006), dass der ANL nicht mit der Sprachverständlichkeit im Störgeräusch zusammenhängt. Dagegen wurde festgestellt, dass der ANL von der Art der Hintergrundgeräusche wie z. B. Stimmengewirr, stationäres Rauschen und Musik abhängig ist (Nabelek *et al.*, 1991; Smiley *et al.*, 2006; Gordon-Hickey und Moore, 2007). Mueller *et al.* (2006) wiesen mit diesem Messverfahren den Nutzen einer digitalen Störgeräuschreduktion in einem Hörgerät nach.

Die subjektiv wahrgenommene Sprachverständlichkeit könnte auch durch Störgeräuschreduktionen beeinflusst werden. Sie kann mit dem Just Follow Conversation Test (JFCT, Larsby und Arlinger (1994); Borg *et al.* (1998); Magnusson *et al.* (2001); Neumann und Jensen (2001)) oder der HörTech Einpegelung der Individuellen Satzverständlichkeitsschwelle (HEISS, Latzel *et al.* (2003)) dokumentiert werden. In beiden Verfahren wird der Versuchsperson ein Sprachsignal zusammen mit einem Hintergrundgeräusch dargeboten. Sie hat die Aufgabe, den Pegel der Sprache (oder auch des Hintergrundgeräusches) zu verändern, bis sie die Sprache gemäß ihrem subjektiven Kriterium gerade eben noch versteht. Ergebnis ist der SNR, der diese Einstellung der Pegel beschreibt. JFCT und HEISS unterscheiden sich im dargebotenen Sprachsignal. Im JFCT wird die kontinuierliche Aufnahme z. B. einer Geschichte verwendet. Im HEISS besteht das Sprachsignal aus Sätzen des Göttinger Satztestes (Kollmeier und Wesselkamp, 1997), die kontinuierlich präsentiert werden.

Ein Maß, das die objektiv gemessene und die individuell wahrgenommene Sprachverständlichkeitsschwelle miteinander vergleicht, ist die Perceptual Performance Discrepan-

cy (PPDIS). Dies ist nach Saunders und Cienkowski (2002) die Differenz zwischen einer individuell wahrgenommenen und der objektiv gemessenen Sprachverständlichkeitsschwelle. Versuchspersonen mit einer positiven PPDIS überschätzen ihre Hörfähigkeit und Versuchspersonen mit einer negativen PPDIS unterschätzen sie. Saunders *et al.* (2004, 2005, 2006) verwendeten zur Ermittlung der Schwellen den Hearing in Noise Test (HINT). Dieser wurde auf eine ähnliche Weise wie der Oldenburger Satztest (OLSA) (Wagener *et al.*, 1999a,b,c) und der JFCT durchgeführt. Saunders *et al.* (2005) beschreiben, dass die PPDIS u. a. signifikant mit dem beschriebenen Nutzen der Hörgeräte im Hintergrundgeräusch und der Tragedauer von Hörgeräten korreliert.

Die vorliegende Studie untersucht die genannten Testverfahren zur Dokumentation des Nutzens von einkanaligen Störgeräuschreduktionen. Der Schwerpunkt liegt dabei auf dem ANLT. Dessen Ergebnisse sollen verglichen werden mit denen des OLSA, des JFCT, der PPDIS und des Paarvergleichs. Verwendung findet ein Messaufbau, wie er z. B. in der Hörgeräteentwicklung genutzt werden könnte. Die Signale werden dabei nur von der Störgeräuschreduktion verarbeitet und direkt ausgegeben. Es werden keine weiteren Komponenten z. B. einer vollständigen Hörgeräteverarbeitung simuliert. Untersucht werden soll, ob der ANLT zur Entwicklung von einkanaligen Störgeräuschreduktionen verwendet werden kann und eindeutigere Ergebnisse als Sprachverständlichkeitstests erreicht, oder weniger aufwendig ist als Paarvergleiche.

4.3 Methode

4.3.1 Versuchspersonen

An den Untersuchungen nahmen zehn normalhörende Versuchspersonen (vier Frauen, sechs Männer) teil. Sie waren zwischen 22 und 41 Jahren alt (mittleres Alter: 28 Jahre) und zeigten höchstens bei einer der Oktavfrequenzen zwischen 0,125 und 8 kHz eine maximale Hörschwelle von 20 dB HL. Alle hatten bereits Erfahrung mit dem ANLT und anderen Hörtests gesammelt. Außerdem wurden Messungen mit elf Schwerhörigen durchgeführt (acht Frauen, drei Männer). Die schwerhörigen Versuchspersonen waren zwischen 13 und 67 Jahren alt (mittleres Alter: 43,3 Jahre). Ihr Hörverlust ist in Abb. 4.1 dargestellt. Sie hatten keine Erfahrung mit dem ANLT und wurden für ihren Aufwand mit 10 Euro/h entschädigt. Bei den Versuchspersonen handelte es sich z. T. um Studenten und Mitarbeiter des Studienganges Hörtechnik und Audiologie in Oldenburg. Diese hatten große Kenntnisse über das OLSA-Sprachmaterial, das sie in verschiedenen Versuchen bereits gehört hatten. Alle anderen Versuchspersonen hatten sich freiwillig für die Teilnahme an Studien im Haus des Hörens in Oldenburg gemeldet und wurden aus der entstandenen Versuchspersonendatei rekrutiert.

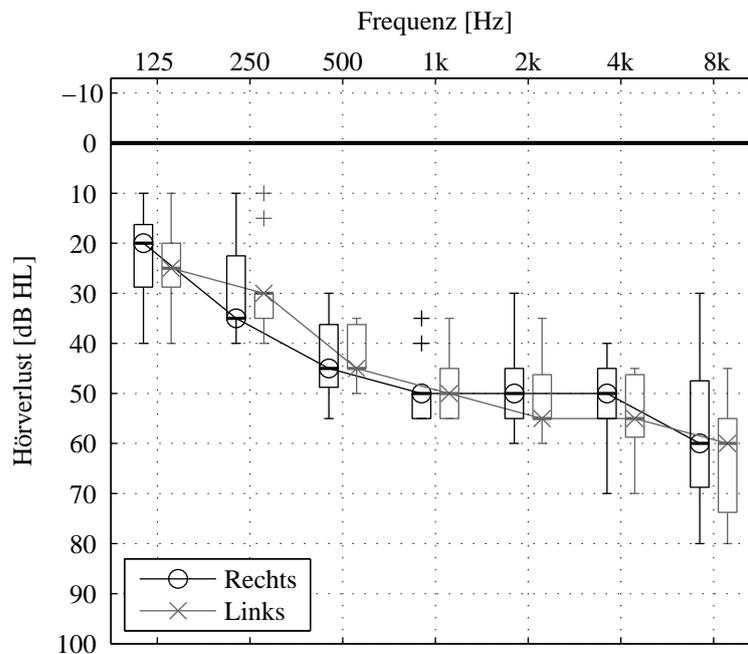


Abb. 4.1: Hörverlust der schwerhörigen Versuchspersonen, gemessen mit einem Tonaudiometer.

4.3.2 Aufbau und Testmaterial

Für alle Messungen wurden die Signale von einem PC mit Matlab (Version 7.3) verarbeitet, über eine Soundkarte (RME AD/DA-Interface ADI 2) an einen Kopfhörerverstärker (Tucker Davis Technologies HB7 Headphonedriver) übertragen und über einen Kopfhörer (Sennheiser HDA 200) in einer Hörkabine (Soundblocker Audiometrikabine, Größe: B) wiedergegeben. Dieser Aufbau wurde freifeldentzerrt und nach DIN EN 60645-2 (1997) und DIN EN ISO 389-8 (2004) kalibriert. Anhang A enthält eine detaillierter Beschreibung des Messaufbaus und der Kalibrierung.

Als Sprachsignal wurden die Sätze des OLSA verwendet (Wagener *et al.*, 1999a,b,c; Wagener, 2003). Die Sätze haben eine einfache syntaktische Struktur (Name, Verb, Zahl, Adjektiv, Objekt), kaum Redundanz und sie ergeben aufgrund ihrer Generierung nicht immer einen Sinn. So wird verhindert, dass sie auswendig gelernt werden können. Im OLSA können 40 Testlisten verwendet werden, wobei eine Testliste aus 30 Sätzen besteht. Eine Testliste wird der Versuchsperson mit Störgeräusch dargeboten und sie muss die verstandenen Sätze wiederholen. Abhängig davon wie viele Wörter eines Satzes verstanden werden, wird der Pegel der Sätze verändert bis ein SRT 50%-igen Wortverstehens erreicht wird. Der Pegel des Rauschens wird konstant gehalten. Alternativ dazu kann auch die Sprache bei konstantem Pegel präsentiert und der Pegel des Hintergrundgeräusches variiert werden.

Für eine kontinuierliche Wiedergabe im JFCT und ANLT wurden die Sätze des OLSA in randomisierter Reihenfolge mit Pausen zwischen 0,25 und 0,6 s dargeboten. Diese Werte liegen in einem Bereich, der von Campione und Veronis (2002) als mittlere Pausenlänge in fließend gesprochener Sprache bezeichnet wird. Als Hintergrundgeräusch

wurde für alle Tests das Rauschen des OLSA gewählt. Es hat das gleiche Spektrum wie die Sprachaufnahmen des OLSA. Diese Signale wurden gewählt, um einen Vergleich der Ergebnisse des ANLT, des OLSA und des JFCT zu ermöglichen.

4.3.3 Einkanalige Störgeräuschreduktionen

Mit den subjektiven Tests wurden drei unterschiedliche einkanalige Störgeräuschreduktionen geprüft. Untersucht wurde ein optimaler Algorithmus, der über ein a priori Wissen des Hintergrundgeräusches verfügt, d. h., dass das Störgeräuschsignal getrennt von der Sprache vorliegt und keine Störgeräuschschätzung benötigt wird. Da Sprachsignal und Hintergrundgeräusch in realen Umgebungen am Eingang der Störgeräuschreduktion ein Gesamtsignal bilden, das nicht mehr getrennt werden kann, ist diese Verarbeitung unrealistisch. Das a priori Wissen wurde in einer Verstärkungsfunktion nach Wiener angewendet (Vary *et al.*, 1998). Dieses Filter ist danach optimal im Sinne Wieners. Seine maximale theoretische Reduktion wurde auf 6 dB begrenzt. D. h., die Signale konnten durch die Verwendung eines Spectral Floor (Kroschel, 2004) in der Störgeräuschreduktion um maximal 6 dB abgesenkt werden. Im Folgenden wird dieser Algorithmus mit *Optimal* bezeichnet.

Außerdem wurden zwei realistische Algorithmen verwendet. Sie schätzen aus dem Gesamtsignal das Rauschen mit Hilfe des Minima Controlled Recursive Averaging (Cohen und Berdugo, 2002) und filtern das Signal nach der Methode der Spektralen Subtraktion der Leistungsdichtespektren (Vary *et al.*, 1998). Ihre maximale theoretische Reduktion wurde auf 6 dB und 8 dB begrenzt. Sie werden im Folgenden mit *Real6dB* und *Real8dB* bezeichnet. Zusätzlich zu den Messungen mit den Störgeräuschreduktionen *Optimal*, *Real6dB* und *Real8dB* wurden Messungen ohne Störgeräuschreduktion durchgeführt. Die Ergebnisse sind mit *NoAlgo* gekennzeichnet. Alle Störgeräuschreduktionen arbeiten mit der Weighted Overlap-Add-Methode, in der eine Fast-Fourier-Transformationslänge von 128 Samples mit 50%-iger Überlappung verwendet wird. Die Signale wurden mit einer Frequenz von 32000 Hz abgetastet.

Die Störgeräuschreduktionen verarbeiteten neben dem Gesamtsignal aus Sprache und Hintergrundgeräusch auch zeitgleich die separaten Sprach- und Rauschsignale mit den für das Gesamtsignal ermittelten Parametern. Dies kann zur exakten Bestimmung des SNR am Ausgang des Systems verwendet werden, da die Filterung linear geschieht. Aufgrund der getrennten Verarbeitung des Sprach- und Rauschsignals wurde die tatsächliche SNR-Verbesserung für jede Störgeräuschreduktion in Abhängigkeit vom SNR am Eingang ermittelt. Sie ist definiert als Differenz zwischen dem SNR in dB der Signale am Aus- und Eingang der Störgeräuschreduktion und wird im Folgenden tatsächliche SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ genannt. Sie ist in Abb. 4.3 dargestellt. Als Sprachsignal wurden 27 willkürlich gewählte Sätze des OLSA, die zusammen einer Länge von 1,02 min entsprachen, kontinuierlich hintereinander dargeboten. Das OLSA-Rauschen wurde als Hintergrundgeräusch verwendet, das 7 s vor der Sprachwiedergabe begann. Durch den früheren Beginn des Rauschens konnte es durch die Störgeräuschreduktion geschätzt werden. Zur Bestimmung des SNR wurden nur die Sprach- und Rauschsi-

gnale analysiert, die innerhalb der 1,02 min langen kontinuierlichen Sprachwiedergabe verarbeitet wurden. Abb. 4.2 zeigt die zeitliche Abfolge der verwendeten Signale. Die in Abb. 4.3 dargestellte tatsächliche SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ erreicht ihr Maximum, wenn der Eingangs-SNR zwischen 2 und 6 dB SNR liegt. Die Maxima liegen bei 3,3 dB SNR (*Optimal*), 1,9 dB SNR (*Real8dB*) und 1,6 dB SNR (*Real6dB*). Die tatsächliche SNR-Verbesserung ist bei allen Störgeräuschreduktionen niedriger als die theoretisch mögliche Reduktion, die nur die maximale Pegelverringerung beschreibt und nicht die tatsächliche SNR-Verbesserung berücksichtigt.

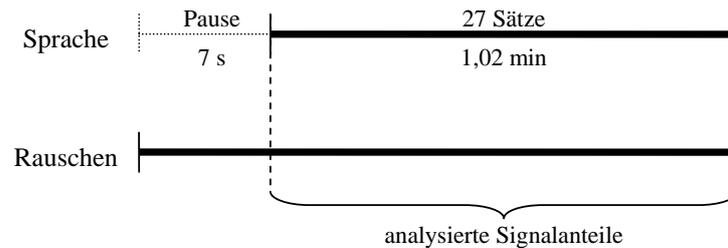


Abb. 4.2: Zur Analyse der tatsächlichen SNR-Verbesserung genutzte Signale.

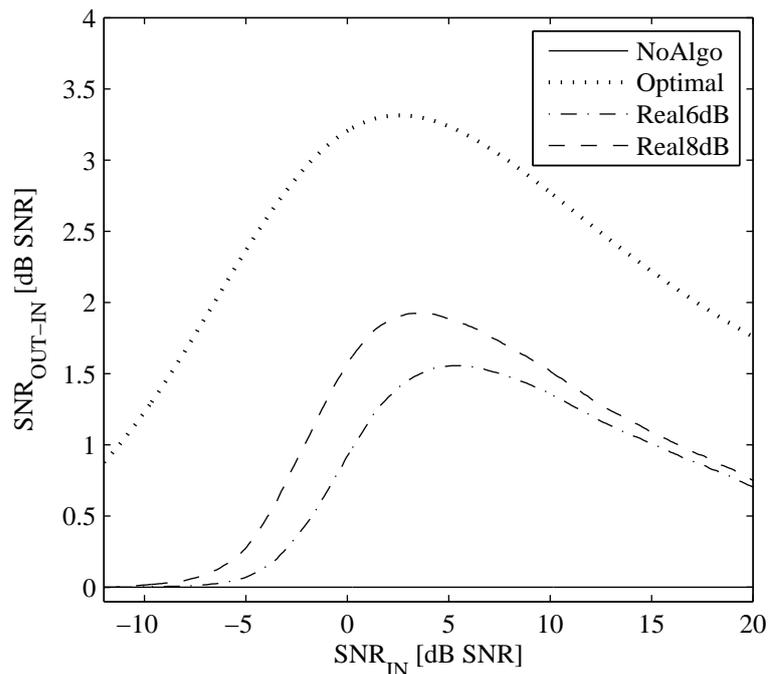


Abb. 4.3: Differenz zwischen dem SNR der Signale am Aus- und Eingang der Störgeräuschreduktion ($\text{SNR}_{\text{OUT-IN}}$), aufgetragen über dem SNR am Eingang (SNR_{IN}).

Da einkanalige Störgeräuschreduktionen auf das gestörte Sprachsignal eine Dämpfung pro Frequenzband anwenden, wird der Pegel der Sprache und des Rauschens verändert, wenn beide Signalanteile sich spektral überlappen. Damit der Pegel der Sprache im ANLT konstant bleibt, ist es notwendig, diesen Pegelverlust auszugleichen. Dazu wurde das getrennt verarbeitete Sprachsignal genutzt. Die Pegel des Sprachsignals vor und nach der Störgeräuschreduktion wurden verglichen und daraus die benötigte Verstär-

kung berechnet. Um eine sprunghafte Veränderung des Gesamtsignalpegels zu verhindern, wurde die berechnete Verstärkung mit einem Fenster der Länge 1 s geglättet und auf das Ausgangssignal der Störgeräuschreduktion angewendet.

4.3.4 Messverfahren

Vor Beginn der Untersuchungen wurde mit jeder Versuchsperson eine Tonaudiometrie durchgeführt. Mit Versuchspersonen, denen das Sprachmaterial nicht in ausreichendem Maße bekannt war, wurde vor Beginn der Messungen geübt. Dazu wurden zufällig drei Testlisten mit 30 Sätzen ausgewählt und bei einem konstanten Pegel ohne Hintergrundgeräusch präsentiert. Gewählt wurde ein Pegel, bei dem die Versuchspersonen die Sätze ohne Anstrengung verstehen konnten. Sie hatten dann die Aufgabe, die Sätze zu wiederholen. Fehler wurden korrigiert, indem den Versuchspersonen die richtige Antwort vorgesprochen wurde. In den anschließenden Messungen wurde zuerst mit Hilfe des ANLT der MCL und dann der BNL bestimmt. Der ermittelte MCL wurde für die nachfolgenden Messungen als Sprachpegel verwendet. Es folgten der JFCT und der OLSA. In einem zweiten Termin wurde der Paarvergleich durchgeführt.

ANLT

Basierend auf den Beschreibungen nach Nabelek *et al.* (1991, 2006) wurde mit jeder Versuchsperson der MCL für Sprache bestimmt. Dazu wurde sie angewiesen, den Pegel der Sprache lauter als angenehm, danach leiser als angenehm und abschließend auf einen angenehmen Pegel (MCL) einzustellen, dem sie über eine längere Zeit zuhören würde. Der Pegel des ersten Satzes betrug 30 dB(A). Zur Einstellung des lauten und leisen Pegels wurden 5 dB-Schritte und zur Einstellung des MCL wurden 2 dB-Schritte verwendet. Diese Einstellungen wurden dreimal vorgenommen und der Median der MCL-Werte ermittelt. Die Instruktion zur Bestimmung des MCL ist in Anhang B.3 dargestellt.

Zur Bestimmung des BNL wurde die Sprache konstant beim Median der MCL-Werte und zusätzlich das Hintergrundgeräusch bei einem Pegel von 30 dB(A) präsentiert. Der Pegel des Hintergrundgeräusches wurde dann durch die Versuchsperson verändert. Zuerst wählte sie eine Lautstärke, bei der sie die Sätze nicht mehr verstehen konnte, danach sollte sie den Pegel verringern, bis die Sätze deutlich zu verstehen waren. Zuletzt wählte sie die maximale Lautstärke des Hintergrundgeräusches, die sie über einen längeren Zeitraum tolerieren würde, ohne sich belästigt zu fühlen. Sie sollte den Sätzen zuhören können, ohne zu ermüden. Dabei sollte sie sich vorstellen, im Radio einer Dokumentation zu folgen und auf den Pegel eines Hintergrundgeräusches (z. B. eines Staubsaugers) direkt Einfluss nehmen zu können. Die Instruktion zur Bestimmung des BNL und ANL ist in Anhang B.4 dargestellt. Auch hier wurde der maximale und minimale Pegel des Rauschens mit einer Schrittgröße von 5 dB und der tolerierbare Pegel mit einer Schrittgröße von 2 dB eingestellt. Aus der Differenz zwischen MCL und BNL wurde der ANL berechnet ($ANL = MCL - BNL$). Für jede der vier Testsituationen (*Optimal*, *Real6dB*, *Real8dB* und *NoAlgo*) wurde der BNL dreimal ermittelt. Die verwendeten Testsituatio-

nen wurden innerhalb der Messung randomisiert. Insgesamt führte die Versuchsperson die Bestimmung des ANL also zwölfmal durch.

Zur Einstellung des Pegels im ANLT wurde den Versuchspersonen eine graphische Oberfläche auf einem Bildschirm dargeboten (s. Anhang B.3). Dort wurde die aktuelle Instruktion (z.B. „...lauter als angenehm...“ oder „...leiser als angenehm...“) präsentiert und die Versuchspersonen hatten die Möglichkeit, mit Hilfe von Schaltflächen die Lautstärke der Signale zu verändern. Die Schaltflächen konnten mit der Maus des PCs ausgewählt werden. Sobald eine PegelEinstellung beendet war, bestätigte die Versuchsperson sie mit einem OK-Button und die Instruktion für den nächsten Messschritt im ANLT wurde präsentiert. Der OK-Button wurde erst 2 s nach der letzten Pegelveränderung auswählbar, damit das Einschwingen der Störgeräuschreduktion abgeschlossen war, bevor die Versuchsperson antwortete.

JFCT

Im JFCT betrug der Pegel der Sprache MCL und der des Hintergrundgeräusches konnte durch Bewegung einer Trackball-Maus kontinuierlich von den Versuchspersonen verändert werden. Sie konnten so SNR-Werte zwischen ± 15 dB SNR einstellen. Ihre Aufgabe war es, einen Pegel des Hintergrundgeräusches zu wählen, bei dem die Sätze gerade eben noch verständlich waren. Die Instruktion ist in Anhang B.5 dargestellt. Hatten die Versuchspersonen diesen Pegel eingestellt, bestätigten sie ihn mit der Leertaste der Tastatur. Die Betätigung der Taste konnte erst erfolgen, nachdem der Rauschpegel 2 s konstant gehalten wurde. So wurde sichergestellt, dass die Störgeräuschreduktion eingeschwungen war. Das Rauschen wurde nach Betätigung der Leertaste randomisiert um 5-10 dB angehoben oder abgesenkt, und die Versuchspersonen stellten erneut ihre subjektive Sprachverständlichkeitsschwelle ein. Diese Einstellungen wurden zehnmal für jede der vier Testsituationen wiederholt. Jede Versuchsperson führte daher insgesamt 40 Bestimmungen durch, die in zwei Blöcke à 20 Messungen unterteilt waren. Die verwendeten Testsituationen wurden innerhalb jeder Messung randomisiert. Die Instruktion ist in Anhang B.5 dargestellt.

OLSA

Mit dem OLSA wurde ein SRT 50%-igen Satzverstehens ermittelt. Eine Instruktion für diesen Test findet sich in Anhang B.6. Die Sätze wurden bei MCL präsentiert und das Rauschen in Abhängigkeit von den verstandenen Sätzen in seinem Pegel verändert. Abweichend vom vorgeschlagenen Verfahren, dass Word-Scoring nutzt (Wagener, 2003), wird in dieser Studie Sentence-Scoring verwendet. Dieses Vorgehen wurde gewählt, um den Störgeräuschreduktionen größere SNR-Verbesserungen zu ermöglichen. Für jede Testsituation wurde die Messung mit einer neuen Testliste durchgeführt. Die Reihenfolge der Testsituationen war randomisiert. Um den einkanaligen Störgeräuschreduktionen die Schätzung des Hintergrundgeräusches zu ermöglichen, wurde das Rauschen kontinuierlich dargeboten und erst 2 s nach einer Pegelveränderung der Satz präsentiert.

Paarvergleich

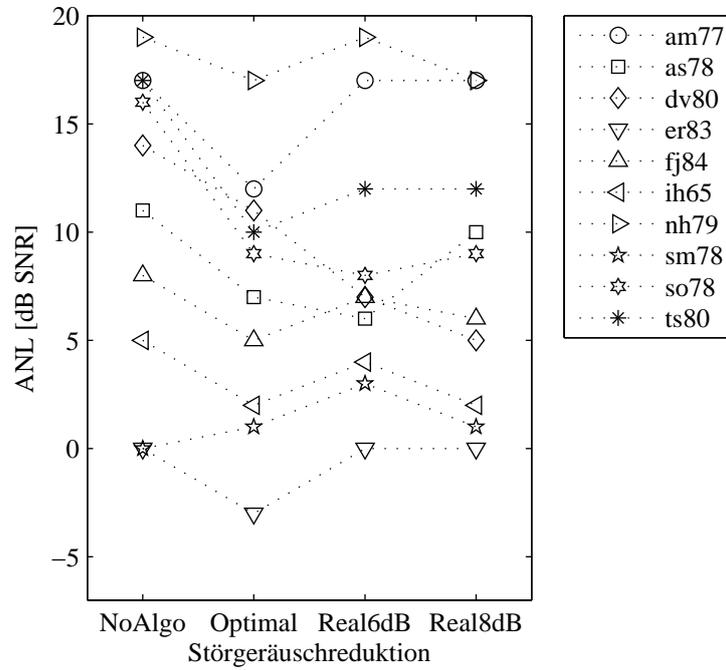
Für den Paarvergleich wurde der Satz „Britta verleiht elf alte Bilder.“ aus dem OLSA zusammen mit dem Hintergrundgeräusch als Klangbeispiel dargeboten. Der Pegel des Satzes betrug MCL. Das Rauschen wurden bei verschiedenen Pegeln verwendet, so dass verschiedene SNR-Werte präsentiert wurden. Diese waren: 1) vier individuelle ANL-Werte, ermittelt in den unterschiedlichen Testsituationen, 2) der mediane ANL aller normalhörenden Versuchspersonen, ermittelt ohne Störgeräuschreduktion (12,5 dB SNR, s. Tab. 4.1), 3) der mediane ANL aller normalhörenden Versuchspersonen mit Störgeräuschreduktion (7,5 dB SNR, s. Tab. 4.1), 4) der maximale und 5) der minimale ANL-Wert, der von Normalhörenden eingestellt wurde (19 dB SNR und -3 dB SNR, s. Tab. 4.1). Die Klangbeispiele wurden mit den Störgeräuschreduktionen verarbeitet. Jedes Klangbeispiel wurde mit jedem anderen bei gleichem SNR im Paarvergleich dargeboten. Die Versuchspersonen hatten das Beispiel zu wählen, dass die bessere Gesamtqualität aufwies. Insgesamt führten die Versuchspersonen 48 Vergleiche (sechs Vergleiche pro SNR) durch. Dazu wurde ihnen eine Bildschirmoberfläche präsentiert, auf der sie die Klangbeispiele anhören und auswählen konnten. Die Instruktion für den Paarvergleich ist in Anhang B.7 dargestellt.

4.4 Ergebnisse

ANLT

Zur Auswertung der Ergebnisse des ANLT wurde aus den drei ANL-Werten je Testsituation der Medianwert errechnet. Die Einzelergebnisse für jede Situation und jede Versuchsperson sind in Anhang E.1 aufgelistet und in Abb. 4.4a für die Normalhörenden und Abb. 4.4b für die Schwerhörigen dargestellt. Die Ergebnisse zeigen, dass durch die Verwendung der optimalen Störgeräuschreduktion in der Regel eine Verringerung des ANL erreicht wurde, d. h., die Versuchspersonen tolerierten einen höheren Pegel des Hintergrundgeräusches mit der Störgeräuschreduktion *Optimal* als ohne Störgeräuschreduktion (*NoAlgo*). Werden die realen Störgeräuschreduktionen (*Real6dB* und *Real8dB*) verwendet, kann kein eindeutiger Trend erkannt werden. Einige Versuchspersonen stellten einen höheren Pegel des Hintergrundgeräusches ein, andere dagegen nicht. Unterschiede zur Situation ohne Störgeräuschreduktion (*NoAlgo*) wurden mit einem Wilcoxon Vorzeichen-Rangtest auf Signifikanz bei einem Niveau von $\alpha = 0,05$ geprüft und für die Normalhörenden mit *Optimal* ($p = 0,004$) und *Real8dB* ($p = 0,023$) ermittelt. Die Ergebnisse der Schwerhörigen zeigen nur für die optimale Störgeräuschreduktion einen signifikanten Unterschied ($p = 0,001$). In Anhang E.2 sind die ermittelten Irrtumswahrscheinlichkeiten aller Vergleiche aufgelistet. Tab. 4.1 zeigt die medianen ANL-Werte bei unterschiedlichen Störgeräuschreduktionen sowie maximale und minimale eingestellte ANL-Werte. Die Werte der Normalhörenden wurden verwendet, um die SNR-Werte für die Darbietung der Signale im Paarvergleich zu bestimmen.

(a) ANL - Normalhörende



(b) ANL - Schwerhörige

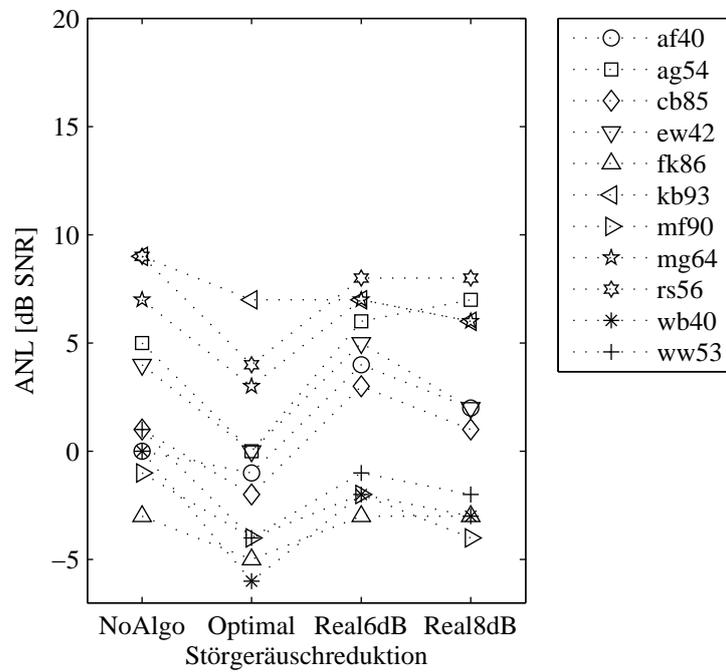
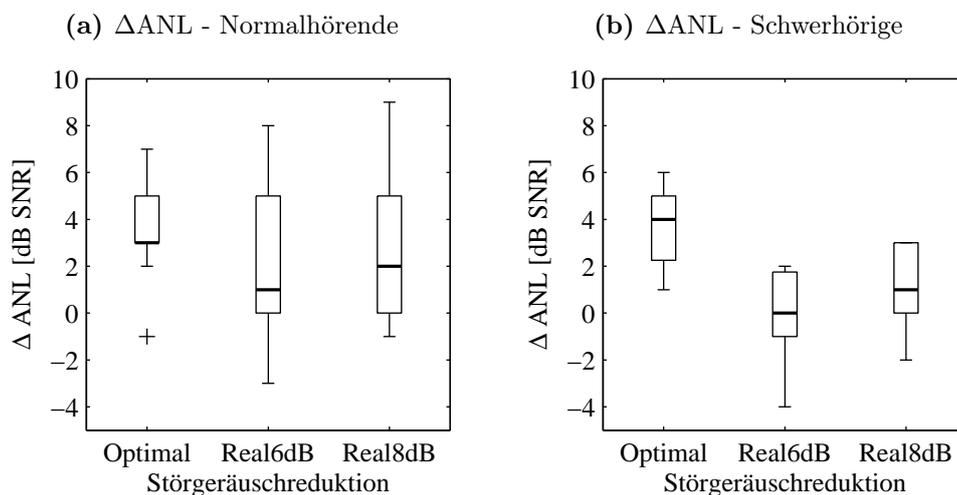


Abb. 4.4: ANL-Werte der (a) Normalhörenden und der (b) Schwerhörigen.

Tab. 4.1: Mediane ANL-Werte in dB SNR der unterschiedlichen Störgeräuschreduktionen und der maximale und minimale eingestellte ANL-Wert.

| | Normalhörende | Schwerhörige |
|---------|---------------|--------------|
| NoAlgo | 12,5 | 1 |
| Optimal | 8 | -1 |
| Real6dB | 7 | 4 |
| Real8dB | 7,5 | 2 |
| Maximum | 19 | 9 |
| Minimum | -3 | -6 |

Damit der Nutzen durch eine Störgeräuschreduktion unabhängig von den interindividuellen Streuungen betrachtet werden kann, wurde die individuelle Differenz zwischen dem medianen ANL in der Situation ohne Störgeräuschreduktion und dem medianen ANL mit Störgeräuschreduktion gebildet. Dieses Maß wird im Folgenden ΔANL genannt ($\Delta\text{ANL} = \text{ANL}_{\text{NoAlgo}} - \text{ANL}_{\text{Algo}}$). Die Ergebnisse sind für die normalhörenden Versuchspersonen in Abb. 4.5a und für die schwerhörigen Versuchspersonen in Abb. 4.5b dargestellt. Ein größerer Nutzen der Störgeräuschreduktion führt zu einem höheren ΔANL . Die Normalhörenden erreichten mit der optimalen Störgeräuschreduktion einen medianen ΔANL von 3 dB SNR, Schwerhörige von 4 dB SNR. Mit realen Störgeräuschreduktionen erreichen die Normalhörenden einen ΔANL von 1 dB SNR (*Real6dB*) bzw. 2 dB SNR (*Real8dB*). Schwerhörige wählten ΔANL -Werte von 0 dB SNR (*Real6dB*) bzw. 1 dB SNR (*Real8dB*). Die maximale tatsächliche SNR-Verbesserung $\Delta\text{SNR}_{\text{OUT-IN}}$ von 3,3 dB SNR (*Optimal*), 1,9 dB SNR (*Real8dB*) und 1,6 dB SNR (*Real6dB*, s. Abb. 4.3) konnte damit annähernd erreicht werden.

**Abb. 4.5:** Nutzen der unterschiedlichen Störgeräuschreduktionen, gemessen im ANLT für (a) Normalhörende und (b) Schwerhörige.

JFCT

Zur Auswertung der Ergebnisse des JFCT wurde der Median der zehn JFC-Werte einer Testsituation gebildet. Ein Vergleich der Ergebnisse mit Störgeräuschreduktion mit den Ergebnissen, gemessen in Situationen ohne Störgeräuschreduktion, mit einem Wilcoxon Vorzeichen-Rangtest auf einem Signifikanzniveau von $\alpha = 0,05$ zeigt, dass Normalhörende signifikante Unterschiede nur mit der optimalen Störgeräuschreduktion erreichen ($p = 0,002$). Auch die Schwerhörigen zeigen nur signifikante Unterschiede für die Störgeräuschreduktion *Optimal* ($p = 0,001$). In Anhang E.2 sind die ermittelten Irrtumswahrscheinlichkeiten aller Vergleiche aufgelistet.

Die Ergebnisse des JFCT wurden ebenfalls von der interindividuellen Streuung befreit, indem die Differenz zwischen den Ergebnissen mit und ohne Störgeräuschreduktion gebildet und so der Nutzen der Störgeräuschreduktion ermittelt wurde. Dieser Wert wird im Folgenden ΔJFC -Wert genannt ($\Delta\text{JFC} = \text{JFC}_{\text{NoAlgo}} - \text{JFC}_{\text{Algo}}$). Abb. 4.6 beinhaltet die Ergebnisse für (a) die Normalhörenden und (b) die Schwerhörigen. Beide Gruppen zeigen lediglich einen Nutzen bei Anwendung der optimalen Störgeräuschreduktion, der bei den Normalhörenden 2,3 dB SNR und bei den Schwerhörigen 3,1 dB SNR beträgt.

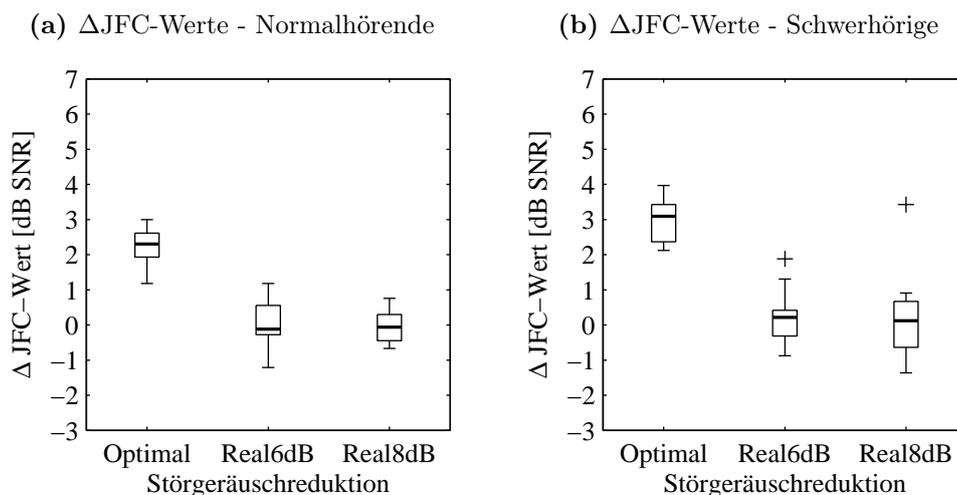


Abb. 4.6: Nutzen der unterschiedlichen Störgeräuschreduktionen, gemessen mit dem JFCT für (a) Normalhörende und (b) Schwerhörige.

OLSA

Die Ergebnisse des OLSA, die SRT-Werte, wurden vergleichbar ausgewertet, wie diejenigen des ANLT und des JFCT. Unterschiede zwischen den SRT-Werten, gemessen ohne und mit den verschiedenen Störgeräuschreduktionen, wurden mit einem Wilcoxon Vorzeichen-Rangtest und einem Signifikanzniveau von $\alpha = 0,05$ geprüft. Auch hier konnten nur signifikante Unterschiede beider Versuchspersonengruppen nur für die optimale Störgeräuschreduktion ermittelt werden (Normalhörende: $p = 0,002$, Schwerhörige: $p = 0,001$). In Anhang E.2 sind die ermittelten Irrtumswahrscheinlichkeiten aller Vergleiche aufgelistet. Außerdem wurde der Nutzen der Störgeräuschreduktion ermittelt.

Er wurde berechnet durch Subtraktion des SRT, gemessen mit Störgeräuschreduktion, vom SRT-Wert, gemessen ohne Störgeräuschreduktion. Dieser Wert wird im Folgenden ΔSRT genannt ($\Delta\text{SRT} = \text{SRT}_{\text{NoAlgo}} - \text{SRT}_{\text{Algo}}$) und ist in Abb. 4.7a für Normalhörende und in Abb. 4.7b für Schwerhörige dargestellt. Die Normalhörenden zeigen einen Nutzen mit der optimalen Störgeräuschreduktion von ca. 1,9 dB SNR, die Schwerhörigen von 3,5 dB SNR.

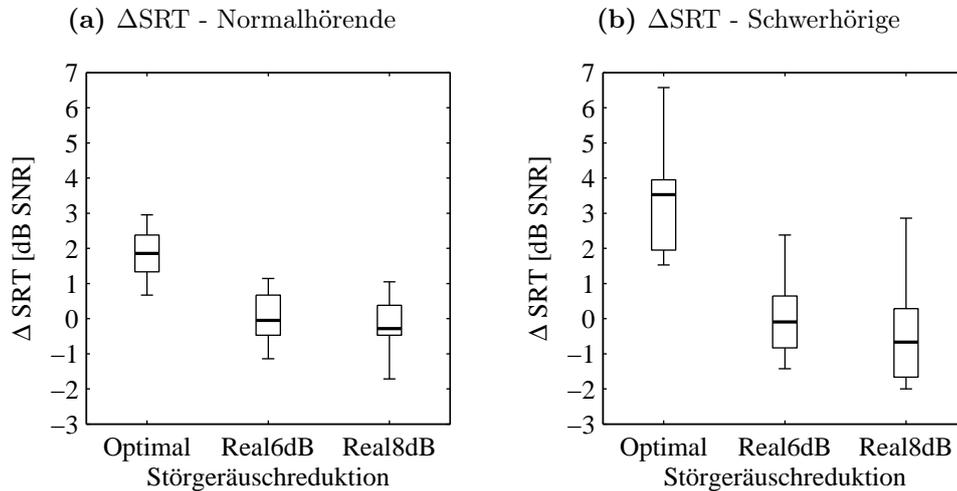


Abb. 4.7: Nutzen der unterschiedlichen Störgeräuschreduktionen, gemessen mit dem OLSA für (a) Normalhörende und (b) Schwerhörige.

PPDIS

Aus der Differenz zwischen dem JFC-Wert und dem SRT wurde die PPDIS berechnet. Tab. 4.2 zeigt die Median-Werte der PPDIS für die unterschiedlichen Testsituationen. Die mediane PPDIS beträgt maximal 1,1 dB SNR. Die JFC-Werte unterscheiden sich also wenig von den SRT-Werten. Die Unterschiede zu den Situationen ohne Störgeräuschreduktion wurden ebenfalls mit einem Wilcoxon Vorzeichen-Rangtest ($\alpha = 0,05$) geprüft. Keines der Ergebnisse ist signifikant unterschiedlich. In Anhang E.2 sind die ermittelten Irrtumswahrscheinlichkeiten aller Vergleiche aufgelistet.

Tab. 4.2: Median der PPDIS in dB SNR für die unterschiedlichen Testsituationen.

| | Normalhörende | Schwerhörige |
|---------|---------------|--------------|
| NoAlgo | 0,952 | 0,082 |
| Optimal | 0,708 | 0,636 |
| Real6dB | 1,050 | 0,437 |
| Real8dB | 1,128 | 0,043 |

Außerdem wurde der Nutzen (ΔPPDIS) durch Subtraktion berechnet ($\Delta\text{PPDIS} = \text{PPDIS}_{\text{NoAlgo}} - \text{PPDIS}_{\text{Algo}}$). ΔPPDIS ist für Normalhörende und Schwerhörige in 4.8 dargestellt. Die Werte der Normalhörenden liegen zwischen 0 dB SNR und 0,3 dB SNR, Werte der Schwerhörigen zwischen 0,2 dB SNR und 0,6 dB SNR.

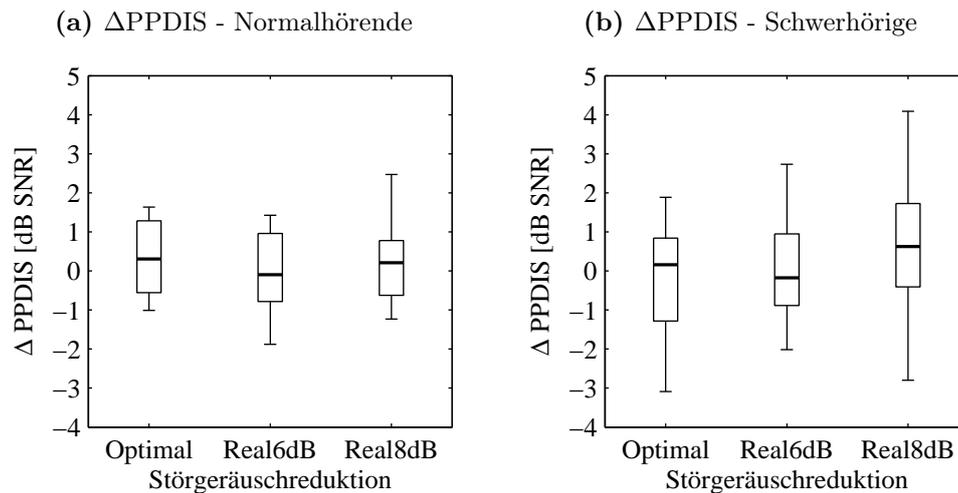


Abb. 4.8: Nutzen der unterschiedlichen Störgeräuschreduktionen ermittelt mit der PPDIS für (a) Normalhörende und der (b) Schwerhörige.

Paarvergleich

Zur Auswertung des Paarvergleiches wurden Häufigkeitsmatrizen für jeden gemessenen SNR erstellt. Dazu wurden die Ergebnisse jeder Versuchsperson in einer 4x4 Matrix dargestellt. Diese enthielt eine 1, wenn die Testsituation der gewählten Zeile i der der Spalte j vorgezogen wurde und 0 für den konträren Fall. Für ein Gesamturteil wurden die Matrizen aller Versuchspersonen zu Häufigkeitsmatrizen addiert. Jede Spalte wurde aufsummiert und durch die Anzahl der mit jeder Störgeräuschreduktion durchgeführten Vergleiche dividiert. Diese Ergebnisse sind als Präferenz in Prozent in Anhang E.3.1 zusammen mit der Konsistenz und der Akkordanz des Paarvergleiches in Anhang E.3.2 dargestellt. Zur Ermittlung eines Gesamtergebnisses wurden die Präferenzen aller SNR-Werte gemeinsam ausgewertet. Abb. 4.9a und Abb. 4.9b zeigt diese Ergebnisse für die Gruppe der Normalhörenden und der Schwerhörigen. Beide Gruppen ziehen Situationen mit optimaler Störgeräuschreduktion denen mit *Real6dB*, *Real8dB* und *NoAlgo* vor. Die Normalhörenden lehnen im Median *Real8dB* deutlicher ab als die Schwerhörigen. Mit einem Wilcoxon Vorzeichen-Rang-Test wurde der signifikante Unterschied der Ergebnisse mit Störgeräuschreduktion und der Ergebnisse ohne Störgeräuschreduktion mit einem Signifikanzniveau von $\alpha = 0,05$ geprüft. Die Präferenzwerte der Normalhörenden sind lediglich für die Algorithmen *Real6dB* und *Optimal* signifikant unterschiedlich (*Optimal*: $p = 0,008$, *Real6dB*: $p = 0,031$). Die Ergebnisse der Schwerhörigen sind für alle drei Störgeräuschreduktionen signifikant unterschiedlich zu *NoAlgo* (für alle gilt: $p = 0,008$). In Anhang E.2 sind die ermittelten Irrtumswahrscheinlichkeiten aller Vergleiche aufgelistet.

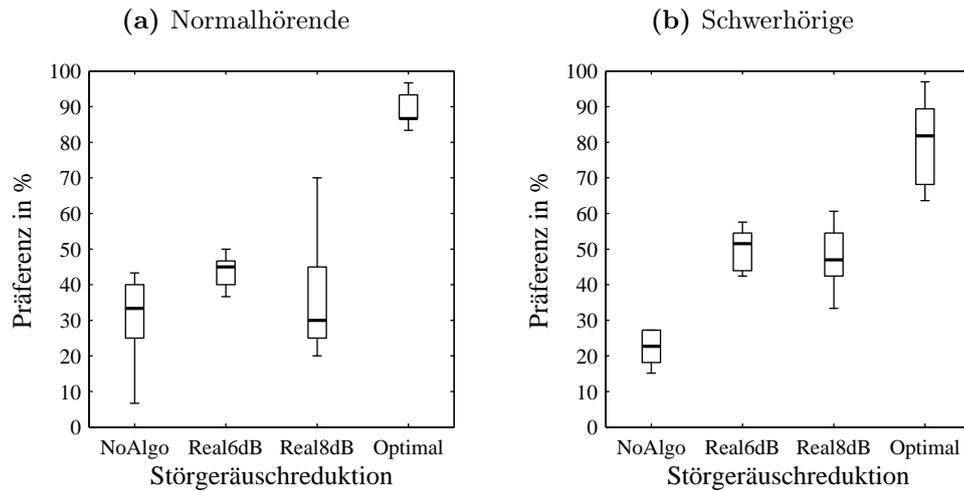


Abb. 4.9: Gesamtergebnis der Paarvergleiche (a) der Normalhörenden und (b) der Schwerhörigen.

Um wie für ANL, JFC-Wert, SRT und PPDIS die Ergebnisse mit und ohne Störgeräuschreduktion zu vergleichen, wurden aus den Paarvergleichen nur die Vergleiche mit *NoAlgo* ausgewählt. Der Anteil der Bevorzungen gegenüber *NoAlgo* für jede Störgeräuschreduktion bei jedem SNR wurde in Prozent bestimmt und in Abb. 4.10 zusammengefasst. Normalhörende Versuchspersonen bevorzugen deutlich die Störgeräuschreduktion *Optimal*. Mit den Algorithmen *Real6dB* und *Real8dB* sinkt die mediane Präferenz auf etwa 50%. Die Präferenz vor allem von *Real8dB* streut sehr stark. Schwerhörige Versuchspersonen favorisieren deutlich alle Störgeräuschreduktionen gegenüber *NoAlgo*.

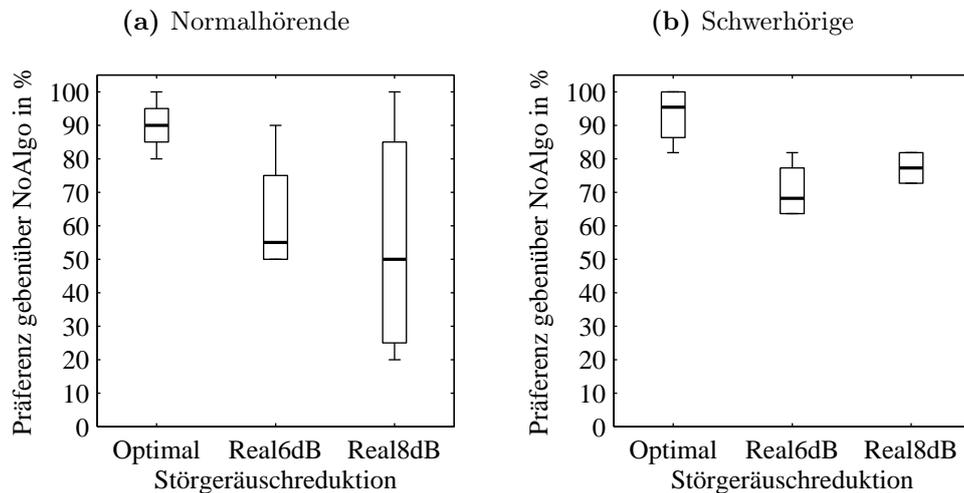


Abb. 4.10: Präferenz gegenüber *NoAlgo*, ermittelt im Paarvergleich. (a) Präferenz der Normalhörenden und (b) der Schwerhörigen.

Vergleich der perzeptiven Maße mit der tatsächlichen SNR-Verbesserung

Um den geringen Nutzen der Störgeräuschreduktionen, ermittelt mit den Sprachverständlichkeitstest JFCT und OLSA (s. Abb. 4.6 und 4.7), zu untersuchen, wurden die gemessenen JFC- und SRT-Werte, die am Eingang der Störgeräuschreduktionen anliegen, der tatsächlichen SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ gegenübergestellt. Abb. 4.11 zeigt im oberen Teil jeder Abbildung die tatsächliche SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ über dem SNR am Eingang der Störgeräuschreduktion SNR_{IN} für jede der verwendeten Störgeräuschreduktionen. Im unteren Teil der Abbildungen sind Boxplots der eingestellten ANL-, JFC- und SRT-Werte der Normalhörenden (schwarz) und Schwerhörigen (grau) dargestellt. Aufgetragen sind für den ANLT der ANL, den JFCT der JFC-Wert und den OLSA der SRT. Die Ergebnisse des JFCT und OLSA wurden bei SNR-Werten am Eingang der Störgeräuschreduktion gemessen, bei denen die realen Störgeräuschreduktionen keinen oder nur einen sehr geringen Nutzen aufweisen (s. Abb. 4.11b und 4.11c). Die gemessenen ANL-Werte überspannen den Bereich, in dem die Störgeräuschreduktionen ihre maximale tatsächliche SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ zeigen.

Um den Zusammenhang zwischen den mit perzeptiven Maßen ermittelten Nutzenwerten der Störgeräuschreduktionen und der tatsächlichen SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ zu untersuchen, wurde der Spearmansche Korrelationskoeffizient r_{sp} berechnet. Seine Irrtumswahrscheinlichkeit p wurde mit einem Permutationstest ermittelt, und mit einem Signifikanzniveau von $\alpha = 0,05$ geprüft. Der Permutationstest untersucht die Hypothese, dass die Korrelation signifikant unterschiedlichen von $r_{\text{sp}} = 0$ ist. Zur Berechnung der Korrelation wurde zu jedem eingestellten perzeptiven Maß (ΔANL , ΔJFC -Wert, ΔSRT , Präferenz gegenüber *NoAlgo*) der $\text{SNR}_{\text{OUT-IN}}$ -Wert bei gleichem SNR am Eingang der Störgeräuschreduktion ermittelt. Für den ΔANL waren die SNR-Werte am Eingang die ANL-Werte, gemessen mit der jeweiligen Störgeräuschreduktion. Für ΔJFC -Wert und ΔSRT wurden die JFC- und SRT-Werte als SNR-Werte am Eingang der Störgeräuschreduktion genutzt. Ein Vergleich der PPDIS mit der tatsächlichen SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ wurde nicht vorgenommen, da die Messwerte keinem eindeutigen SNR am Eingang der Störgeräuschreduktion zugeordnet werden können. Für die Gegenüberstellung mit den Werten des Paarvergleiches wurden nur die Ergebnisse, ermittelt bei den festen SNR-Werten von -3, 7,5, 12,5 und 19 dB SNR, verwendet. Keine der untersuchten Korrelationen ist signifikant. Eine grafische Gegenüberstellung der tatsächlichen SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ mit den perzeptiven Maßen und die dazu ermittelten Korrelationskoeffizienten und Irrtumswahrscheinlichkeiten sind in Anhang E.4 dargestellt.

Vergleich der ANL-Werte mit den anderen perzeptiven Maßen

Zur Untersuchung des Zusammenhanges der perzeptiven Messverfahren mit dem ANL wurden die ANL-Werte jeder Versuchsperson und jeder Testsituation mit Hilfe einer Korrelation mit den Ergebnissen des JFCT, des OLSA, der PPDIS und des Paarvergleiches verglichen. Dazu wurde der Spearmansche Rangkorrelationskoeffizient r_{sp} und seine Irrtumswahrscheinlichkeit p mit einem Permutationstest berechnet und in Tab. 4.3 dargestellt. Für die Signifikanzprüfung wurde ein Niveau von $\alpha = 0,05$ angenommen.

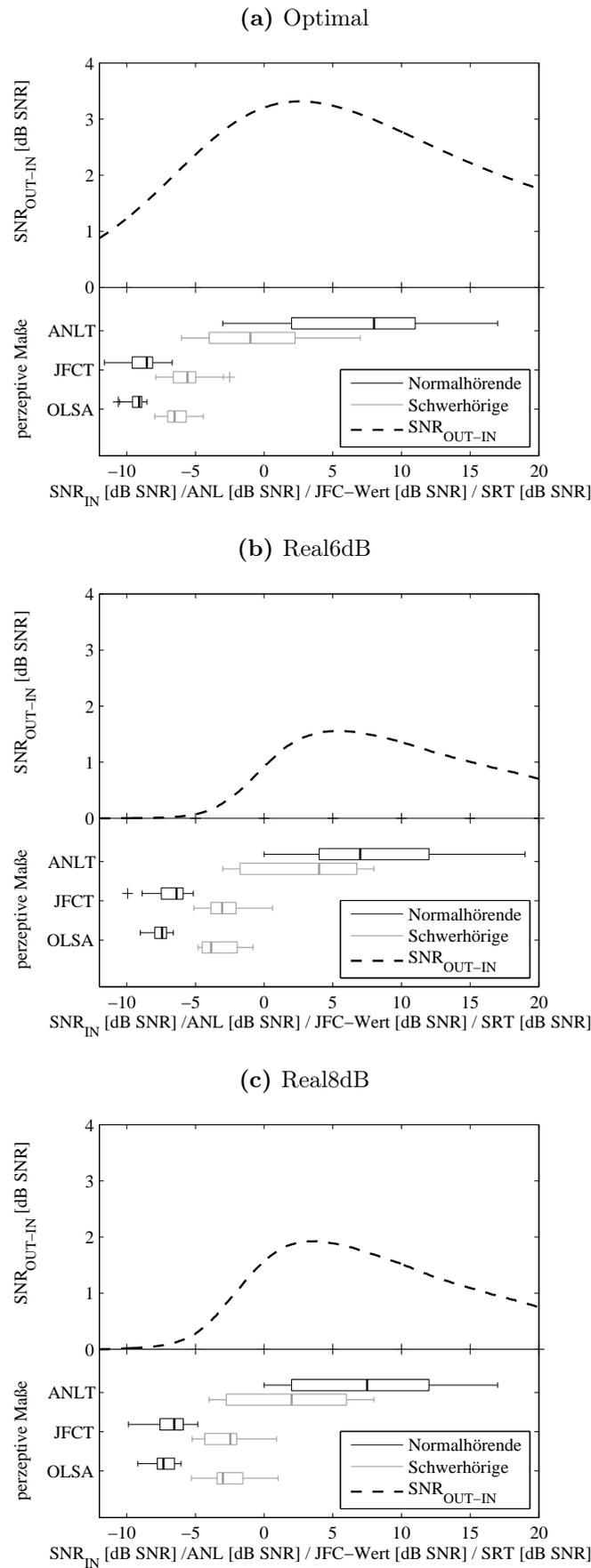


Abb. 4.11: Vergleich der ANL-, JFC- und SRT-Werte mit der tatsächlichen SNR-Verbesserung SNR_{OUT-IN} der Störgeräuschreduktionen (a) *Optimal*, (b) *Real6dB* und (c) *Real 8dB*. Der obere Teil jeder Abbildung zeigt die tatsächliche SNR-Verbesserung SNR_{OUT-IN} über dem SNR am Eingang der Störgeräuschreduktion SNR_{IN} . Im unteren Teil sind für den ANLT die ermittelten ANL-Werte, für den JFCT die JFC-Werte und für den OLSA die SRT-Werte aufgetragen.

Die Werte der Normalhörenden zeigen nur für die Gegenüberstellung der ANL- und JFC-Werte eine signifikante Korrelation, die bei $r_{sp} = -0,38$ liegt. Die Ergebnisse der Schwerhörigen erreichen signifikante Korrelationen zwischen $r_{sp} = 0,58$ und $r_{sp} = 0,74$. Die Ergebnisse des OLSA zeigen einen ähnlichen Zusammenhang wie die des JFCT. Die ermittelten Korrelationen der Normalhörenden sind nicht signifikant unterschiedlich von $r_{sp} = 0$. Die Korrelationen, ermittelt für die Schwerhörigen, sind signifikant und liegen bei $r_{sp} = 0,44$ und $r_{sp} = 0,48$. Der Vergleich der ANLT-Ergebnisse mit den PPDIS-Werten zeigt nur für die Δ ANL- und Δ PPDIS-Werte der Normalhörenden eine signifikante Korrelation von $r_{sp} = 0,41$.

Zum Vergleich der Ergebnisse des ANLT und des Paarvergleiches wurden, wie zuvor bei der Darstellung der Ergebnisse des Paarvergleiches, Häufigkeitsmatrizen erstellt. Allerdings wurden dabei alle Paarvergleiche einer Versuchsperson verwendet. Jede Spalte der Häufigkeitsmatrizen wurde summiert, durch die Gesamtanzahl der Vergleiche dividiert und die Ergebnisse als Prozentwerte aufgetragen. Die ermittelten Präferenzen wurden für jede Versuchsperson und Testsituation zusammen mit den ANL-Werten dargestellt. Alle Ergebnisse zeigen keine signifikante Korrelation. Zum weiteren Vergleich der Ergebnisse des Paarvergleiches mit dem ANLT wurde auch die Präferenz gegenüber *NoAlgo* verwendet. Lediglich für die Ergebnisse Δ ANL und Präferenz gegenüber *NoAlgo* der Schwerhörigen zeigt sich eine signifikante Korrelation von $r_{sp} = 0,44$.

In Anhang E.5 sind die ANLT-Ergebnisse den anderen perzeptiven Maßen in Scatterplots gegenübergestellt und die Ergebnisse einer linearen Regression dokumentiert.

Tab. 4.3: Spearmanscher Korrelationskoeffizient r_{sp} und Irrtumswahrscheinlichkeit p für den Zusammenhang zwischen den Ergebnissen des ANLT und denen des JFCT, des OLSA, der PPDIS und des Paarvergleiches, ermittelt mit unterschiedlichen Störgeräuschreduktionen.

| Untersucher Vergleich | Normalhörende | | Schwerhörige | |
|--|---------------|-------|--------------|-------|
| | r_{sp} | p | r_{sp} | p |
| ANL - JFC | -0,377 | 0,017 | 0,584 | 0,000 |
| Δ ANL - Δ JFC | 0,324 | 0,081 | 0,736 | 0,000 |
| ANL - SRT | -0,025 | 0,878 | 0,481 | 0,001 |
| Δ ANL - Δ SRT | -0,069 | 0,716 | 0,436 | 0,011 |
| ANL - PPDIS | -0,300 | 0,06 | 0,034 | 0,827 |
| Δ ANL - Δ PPDIS | 0,411 | 0,024 | 0,127 | 0,483 |
| ANL - Präferenz | -0,025 | 0,881 | -0,227 | 0,139 |
| Δ ANL - Präferenz gegenüber <i>NoAlgo</i> | -0,135 | 0,478 | 0,435 | 0,011 |

4.5 Diskussion

Ziel dieser Untersuchung war, das Verfahren des ANLT als Untersuchungsform für Störgeräuschreduktionen zu prüfen. Die ANL-Ergebnisse zeigen, dass mit den Störgeräuschreduktionen keine Werte erreicht wurden, die der theoretischen Reduktion entsprachen. Dies wurde auch nicht erwartet, da die tatsächliche SNR-Verbesserung durch die Störgeräuschreduktionen geringer ist. Sie konnte z. T. mit dem Gruppenmittel der ANL-

Ergebnisse nachgebildet werden. Der größte Nutzen wurde durch den optimalen Algorithmus erreicht. Dies legt nahe, dass der ANLT grundsätzlich als Methode zur Untersuchung von Störgeräuschreduktionen geeignet ist. *Real6dB* und *Real8dB* zeigen eine geringe Verbesserung bzw. eine Tendenz, dass sich die Ergebnisse verbessern. Dabei sind die Ergebnisse der Schwerhörigen geringgradig schlechter als die der Normalhörenden. Damit konnten die ANL-Ergebnisse von Mueller *et al.* (2006) nicht in ihrer Deutlichkeit bestätigt werden. Ursächlich für diese Ergebnisse könnten die gemessenen ANL-Werte der Versuchspersonen sein, die eine große Streuung und keine signifikante Korrelation mit der tatsächlichen SNR-Verbesserung aufweisen.

Die JFC- und OLSA-Messungen zeigen die Problematik, dass sie mit SNR-Werten durchgeführt werden, bei denen reale Algorithmen keine SNR-Verbesserung erreichen. Der optimale Algorithmus zeigt in diesen Ergebnissen einen Nutzen, da er das Störgeräusch nicht schätzen muss. Diese Messverfahren können daher nicht zur Untersuchung von realen Störgeräuschreduktionen empfohlen werden.

Die PPDIS, die Differenz zwischen SRT- und JFC-Werten, zeigt keine Verbesserung. Die Versuchspersonen stellen keinen signifikanten Unterschied zwischen den Werten ein. Sie wählen subjektive JFC-Werte, die nahe den SRT-Werten liegen. Eine Beeinflussung dieser Wahl durch die Störgeräuschreduktionen konnte nicht nachgewiesen werden. Ursächlich dafür sind die Einzelwerte, JFC-Wert und SRT. Wie bereits beschrieben, erreichen sie SNR-Werte, bei denen keine SNR-Verbesserung mehr möglich ist. Auch dieses Maß kann deshalb nicht zur Untersuchung von Störgeräuschreduktionen empfohlen werden.

Im Paarvergleich zeigt sich die Bevorzugung der optimalen Störgeräuschreduktion. Außerdem wird die Abhängigkeit des empfundenen Nutzens der Störgeräuschreduktion vom dargebotenen SNR deutlich. Der reale Algorithmus mit einer maximalen Reduktion von 8 dB erzeugt bei Signalen mit einem niedrigen SNR Artefakte, die von den Normalhörenden wahrgenommen werden. Sie lehnen bei niedrigen SNR-Werten diesen Algorithmus sogar gegenüber *NoAlgo* ab. Schwerhörige Versuchspersonen favorisieren *Real8dB* gegenüber *NoAlgo*. Damit wird im Paarvergleich nicht nur die SNR-Verbesserung durch die Störgeräuschreduktion berücksichtigt. Die Versuchspersonen bewerten auch die Klangqualität. Schwerhörige berücksichtigen jedoch in einem größeren Maße die SNR-Verbesserung durch die Störgeräuschreduktion als die Klangqualität, da sie alle Störgeräuschreduktionen gegenüber *NoAlgo* bevorzugen. Die SNR-Verbesserung hat für Schwerhörige vermutlich eine subjektive Verbesserung der Sprachverständlichkeit zur Folge, die favorisiert wird.

Der Vergleich der Ergebnisse des ANLT mit denen des JFCT, OLSA und des Paarvergleichs zeigt für die Normalhörenden keinen positiven signifikanten Zusammenhang. Eine geringe negative Korrelation wurde für die Messwerte ANL und JFC ermittelt. Die aber mit den Δ ANL, Δ JFC und den SRT-Werten nicht bestätigt werden kann. Die perzeptiven Messverfahren untersuchen deshalb also bei den Normalhörenden wahrscheinlich unterschiedliche Aspekte von Sprache im Störgeräusch. Schwerhörige zeigen bei der Untersuchung von Störgeräuschreduktionen einen Zusammenhang zwischen den Messwerten. Je größer der eingestellte ANL- oder Δ ANL-Wert ist, desto größer sind auch SRT- und Δ SRT-Werte und vor allem JFC- und Δ JFC-Werte. Für schwerhö-

rige Versuchspersonen hängt also bei der Untersuchung der Störgeräuschreduktionen der ANL auch z. T. vom Sprachverstehen ab. Die Ergebnisse der Schwerhörigen zeigen außerdem einen Zusammenhang zwischen der Präferenz gegenüber *NoAlgo* und dem Δ ANL. Tendenziell gilt: Je größer der beschriebene Nutzen im ANL ist, desto höher ist auch die Befürwortung der Störgeräuschreduktion im Paarvergleich. Für die schwerhörigen Versuchspersonen ist also die Toleranz des Hintergrundgeräusches z. B. von der Signalqualität und dem Sprachverstehen abhängig. In erster Linie wird ein Schwerhöriger dabei wahrscheinlich das Sprachverstehen berücksichtigen und dann erst die Klangqualität beurteilen. Ihm ist sicherlich wichtiger, die Sprache zu verstehen, als einen unverfälschten Klang der Sprache zu hören.

Diese Ergebnisse zeigen, dass der ANLT grundsätzlich zur Untersuchung von Störgeräuschreduktionen verwendet werden kann. Er zeigt im Gruppenmittel Ergebnisse und Tendenzen, die einen Nutzen von Störgeräuschreduktionen beschreiben können. Schwierigkeiten entstehen durch die interindividuellen Unterschiede der ANL-Werte, wodurch unterschiedliche SNR-Werte am Eingang der Störgeräuschreduktion anliegen. Störgeräuschreduktionen arbeiten bei niedrigen und sehr hohen SNR-Werten weniger effizient als bei mittleren. Es entsteht also eine ungewollte Abhängigkeit zwischen der individuellen Akzeptanz von Hintergrundgeräuschen und der dabei messbaren SNR-Verbesserung. Außerdem zeigen die Werte einen großen Streubereich bzw. große Abweichungen von der möglichen SNR-Verbesserung. Die Streuung ist bei den Ergebnissen des JFCT oder OLSA kleiner. Die Abweichungen könnten durch die verwendete Schrittgröße hervorgerufen sein, die eine Einstellung der z. T. geringen Unterschiede zwischen den Störgeräuschreduktionen nicht möglich macht. Außerdem ist das im ANLT gewählte Kriterium des „tolerierbaren“ Pegels vermutlich keine so eindeutige Schwelle, wie die im JFCT gemessene Sprachverständlichkeitsschwelle. Es ist möglich, dass dieses Kriterium nicht klar durch die Versuchspersonen definiert werden kann. Dennoch hat der ANLT den Vorteil, dass in der Regel bei SNR-Werten gemessen wird, in denen Störgeräuschreduktionen eine SNR-Verbesserung erreichen.

Eine optimale Methode zur Untersuchung realer Störgeräuschreduktionen sollte die Messung bei einem konstanten und positiven SNR durchführen, bei dem eine maximale SNR-Verbesserung erreicht werden kann. Sie sollte ein einfaches Kriterium besitzen und ausreichend feine Schrittgrößen verwenden, mit denen die Unterschiede eingestellt werden können. Eine mögliche Alternative ist evtl. das von Wittkop (1994, 2001) vorgeschlagene Verfahren. Er führte einen Vergleich zwischen einem durch Störgeräuschreduktionen verarbeiteten Referenzsignal und einem veränderlichen unverarbeiteten Signal durch. Die Versuchspersonen hatten die Aufgabe, den SNR des Referenzsignals mit dem unverarbeiteten Signal zu reproduzieren. Diese Methode sollte in einer weiterführenden Untersuchung überprüft werden.

4.6 Schlussfolgerungen

Diese Studie lässt die folgenden Schlüsse zu:

1. Zur Untersuchung von einkanaligen Störgeräuschreduktionen kann der ANLT generell verwendet werden. Probleme bereitet die große Streuung der ANL-Werte. Verschiedene Versuchspersonen stellen unterschiedliche ANL-Werte und damit verschiedene SNR-Werte am Eingang der Störgeräuschreduktion ein. Die Störgeräuschreduktion erreicht bei unterschiedlichen SNR-Werten am Eingang unterschiedliche SNR-Verbesserungen. Außerdem sind die geringen Unterschiede der Störgeräuschreduktionen problematisch. Sie sind z. T. geringer als die verwendete Schrittgröße.
2. OLSA, JFCT und PPDIS sind zur Untersuchung von einkanaligen Störgeräuschreduktionen nicht geeignet. Sie werden bei SNR-Werten gemessen, bei denen in der Regel die Störgeräuschreduktion keine SNR-Verbesserung erreichen kann.
3. Paarvergleiche können zur Untersuchung der Veränderung der Klangqualität durch Störgeräuschreduktionen verwendet werden.
4. Die Ergebnisse der Normalhörenden zeigen nur für ANL und JFC einen geringen Zusammenhang. Diese Werte können mit den weiteren Ergebnissen der Sprachverständlichkeitstests und des Paarvergleichs nicht bestätigt werden. Sie scheinen unterschiedliche Aspekte von Sprache im Störgeräusch zu untersuchen.
5. Die Ergebnisse des ANLT für Schwerhörige korrelieren in geringem Maße mit den Ergebnissen der Sprachverständlichkeitstests und der Präferenz gegenüber *NoAlgo* aus dem Paarvergleich. Der ANL ist damit vermutlich u. a. abhängig vom Sprachverstehen und der Klangqualität.

Kapitel 5

Untersuchung von einkanaligen Störgeräuschreduktionen mit dem Acceptable Noise Level Test und zwei technischen Maßen

5.1 Zusammenfassung

Diese Studie untersucht einkanalige Störgeräuschreduktionen mit dem perceptiven Maß Acceptable Noise Level (ANL) und den technischen Maßen Δ SNR und Modulationstransferfunktion (MTF). Betrachtet wurden drei einkanalige Störgeräuschreduktionen. Ein Algorithmus verwendete ein a priori Wissen des Hintergrundgeräusches und eine Verstärkungsfunktion nach Wiener und seine maximale theoretische Reduktion wurde auf 6 dB SNR beschränkt. Zwei weitere Algorithmen schätzten das Rauschen mit Hilfe des Minima Controlled Recursive Averaging und filterten das Signal nach der Methode der Spektralen Subtraktion der Leistungsdichtespektren. Ihre maximale theoretische Reduktion wurde auf 6 bzw. 8 dB SNR begrenzt. Die Ergebnisse des Acceptable Noise Level Tests (ANLT) zeigten für zehn Normalhörende und elf Schwerhörige, dass der ANLT zur Untersuchung von Störgeräuschreduktionen verwendet werden kann. Problematisch ist die interindividuelle Streuung des ANL. Verschiedene Versuchspersonen, die unterschiedliche ANL-Werte einstellen, erreichen damit verschiedene SNR-Werte am Eingang der Störgeräuschreduktion. Diese reagiert mit unterschiedlichen SNR-Verbesserungen. Δ SNR und MTF eignen sich nur bedingt, die tatsächliche SNR-Verbesserung durch die Störgeräuschreduktion zu beschreiben. Die technischen Maße korrelieren nur vereinzelt mit den Ergebnissen des ANLT.

5.2 Einleitung

Ziel einkanaliger Störgeräuschreduktionen ist eine Verbesserung der Sprachverständlichkeit und/oder der Klangqualität durch Absenkung des Störsignals. Daher werden sie oft bei der Verarbeitung von Sprache im Hintergrundgeräusch eingesetzt und z. B. in Hörgeräten verwendet (Bentler und Chiou, 2006). Einkanalige Störgeräuschreduktionen bestehen in der Regel aus einer Störgeräuschschätzung und einer Verstärkungsregel. Die Störgeräuschschätzung stellt im gestörten Sprachsignal den Hintergrundgeräuschanteil fest. Eine Möglichkeit der Schätzung ist das „Minima Controlled Recursive Averaging“ von Cohen und Berdugo (2002). Es schätzt anhand der Minima des Signals in Sprachpausen das Rauschen. In einem zweiten Schritt wird die frequenzabhängige Verstärkung berechnet. Berechnungsformeln dafür sind z. B. das Wiener Filter oder die Spektrale Subtraktion (Vary *et al.*, 1998). Diese Verstärkungsregeln reduzieren den Pegel in verschiedenen Frequenzbändern, die vom Hintergrundgeräusch dominiert werden.

Neben der Klangqualität und der Sprachverständlichkeit nehmen einkanalige Störgeräuschreduktionen vermutlich Einfluss auf die Akzeptanz von Hintergrundgeräuschen. Zur Untersuchung kann der Acceptable Noise Level Test (ANLT) (Nabelek *et al.*, 1991, 2006) verwendet werden. Dieses Verfahren stellt den akzeptierbaren SNR für Sprache im Hintergrundgeräusch fest. Der so ermittelte SNR wird Acceptable Noise Level (ANL) genannt. Er ist definiert als Differenz zwischen dem Most Comfortable Level (MCL) der Sprache und dem höchsten Pegel des Hintergrundgeräusches, den die Versuchsperson über längere Zeit beim Hören von Sprache akzeptieren würde. Dieser Pegel des Hintergrundgeräusches wird Background Noise Level (BNL) genannt. Versuchspersonen mit einem niedrigen ANL tolerieren höhere Hintergrundgeräuschpegel als Versuchspersonen mit einem hohen ANL. Verschiedene Studien zeigten, dass der ANL unabhängig vom Grad des Hörverlustes, vom Alter, vom Geschlecht der Versuchsperson und von deren Nationalität (englisch, koreanisch) ist und auch davon, ob die Messungen mit oder ohne Hörgerät durchgeführt wurden (Nabelek *et al.*, 1991; Rogers *et al.*, 2003; Nabelek *et al.*, 2004; von Hapsburg und Bangh, 2006; Nabelek *et al.*, 2006). Der ANL war ebenso für die meisten Versuchspersonen nicht abhängig davon, ob monaural oder binaural mit Hörgeräten versorgt wurde (Freyaldenhoven *et al.*, 2006d). Außerdem zeigten Nabelek *et al.* (2004, 2006), dass der ANL nicht mit der Sprachverständlichkeit im Störgeräusch zusammenhängt. Dagegen wurde festgestellt, dass der ANL von der Art der Hintergrundgeräusche wie z. B. Stimmengewirr, stationäres Rauschen und Musik abhängig ist (Nabelek *et al.*, 1991; Smiley *et al.*, 2006; Gordon-Hickey und Moore, 2007). Mueller *et al.* (2006) wiesen mit diesem Messverfahren den Nutzen einer digitalen Störgeräuschreduktion in einem Hörgerät nach.

Zur technischen Untersuchung von Störgeräuschreduktionen eignen sich Modelle und Maße, die die SNR-Veränderung untersuchen. Dazu wurden z. B. der Δ SNR nach Hagerman und Olofsson (2004) oder die Modulationstransferfunktion (MTF) nach Houtgast und Steeneken (1985) und Drullmann *et al.* (1994) vorgeschlagen. Der Δ SNR ist ein Maß für die Veränderung des SNR durch ein Übertragungssystem. Durch Messungen mit Rauschen und invertiertem Rauschen werden die Signalanteile Sprache und Rau-

schen am Ausgang des genutzten Übertragungssystems getrennt. Die SNR-Veränderung kann damit bestimmt werden.

Da durch eine Störgeräuschreduktion der SNR des gestörten Signals erhöht werden soll, erhöht sich damit die Modulationstiefe. Houtgast und Steeneken (1985) stellten die MTF als ein Maß vor, mit dem die Veränderung der Modulationstiefe durch ein Übertragungssystem beschrieben werden kann. Für beliebig modulierte komplexe Eingangssignale kann nach Drullmann *et al.* (1994) die MTF mit Hilfe der Hilberttransformation ermittelt werden. Ist dieser Wert der MTF gleich eins, hat sich die Modulationstiefe nicht verändert. Ist er größer als eins, hat sie sich vergrößert, und ist er kleiner als eins, hat sich die Modulationstiefe verkleinert. Holube *et al.* (2005) untersuchten die Störgeräuschreduktionen verschiedener Hörgeräte mit der MTF. Die meisten Hörgeräte zeigten eine Verringerung des Faktors $\frac{1}{MTF_{PL}}$, was einer Beschreibung der Zunahme der Modulationstiefe durch die Übertragung im Hörgerät entspricht. Fredelake (2006) wies den Nutzen einer Störgeräuschreduktion nicht nur mit der MTF sondern auch mit dem Δ SNR nach. Die hier vorliegende Studie untersucht den Zusammenhang zwischen dem ANL und den technischen Maßen MTF und Δ SNR.

5.3 Methode

5.3.1 Versuchspersonen

An den Untersuchungen nahmen zehn normalhörende Versuchspersonen (vier Frauen, sechs Männer) teil. Sie waren zwischen 22 und 41 Jahren alt (mittleres Alter: 28 Jahre) und zeigten höchstens bei einer der Oktavfrequenzen zwischen 0,125 und 8 kHz eine maximale Hörschwelle von 20 dB HL. Alle hatten bereits Erfahrung mit dem ANLT und anderen Hörtests gesammelt. Außerdem wurden Messungen mit elf Schwerhörigen durchgeführt (acht Frauen, drei Männer). Die schwerhörigen Versuchspersonen waren zwischen 13 und 67 Jahren alt (mittleres Alter: 43,3 Jahre). Ihr Hörverlust ist in Abb. 5.1 dargestellt. Sie hatten keine Erfahrung mit dem ANLT und wurden für ihren Aufwand mit 10 Euro/h entschädigt. Bei den Versuchspersonen handelte es sich z. T. um Studierende und Mitarbeiter des Studienganges Hörtechnik und Audiologie in Oldenburg. Sie hatten große Kenntnisse über das Sprachmaterial des Oldenburger Satztests (OLSA), das sie in verschiedenen Versuchen bereits gehört hatten. Alle anderen Versuchspersonen hatten sich freiwillig für die Teilnahme an Studien im Haus des Hörens in Oldenburg gemeldet und wurden aus der entstandenen Versuchspersonendatei rekrutiert.

5.3.2 Aufbau und Testmaterial

Für alle Messungen wurden die Signale von einem PC mit Matlab (Version 7.3) verarbeitet, über eine Soundkarte (RME AD/DA-Interface ADI 2) an einen Kopfhörerverstärker (Tucker Davis Technologies HB7 Headphonedriver) übertragen und über einen Kopfhörer (Sennheiser HDA 200) in einer Hörkabine (Soundblocker Audiometrikabine, Größe: B) wiedergegeben. Dieser Aufbau wurde freifeldentzerrt und nach DIN EN 60645-2

(1997) und DIN EN ISO 389-8 (2004) kalibriert. Anhang A enthält eine detaillierter Beschreibung des Messaufbaus und der Kalibrierung.

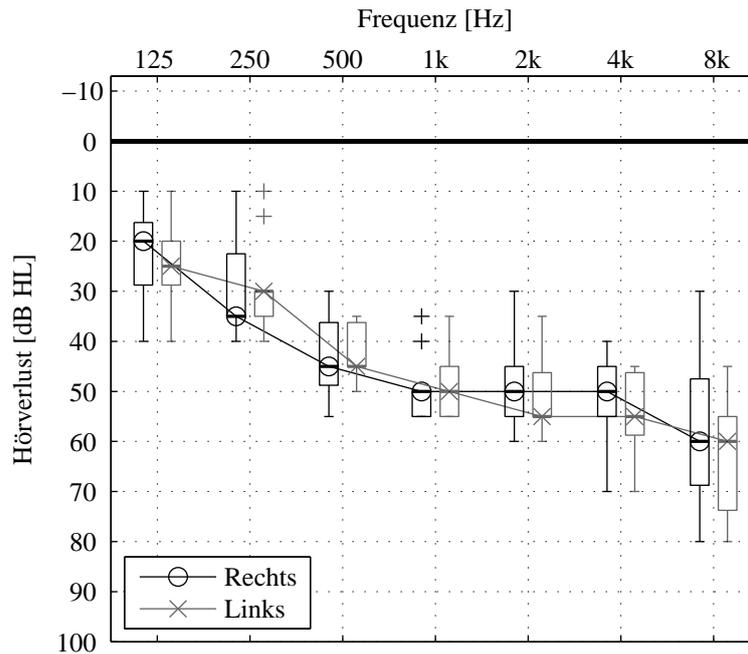


Abb. 5.1: Hörverlust der schwerhörigen Versuchspersonen, gemessen mit einem Tonaudiometer.

Als Sprachsignal wurden die Sätze des OLSA verwendet (Wagner *et al.*, 1999a,b,c; Wagner, 2003). Die Sätze haben eine einfache syntaktische Struktur (Name, Verb, Zahl, Adjektiv, Objekt), kaum Redundanz und sie ergeben aufgrund ihrer Generierung nicht immer einen Sinn. So wird verhindert, dass sie auswendig gelernt werden können. Im OLSA können 40 Testlisten verwendet werden, wobei eine Testliste aus 30 Sätzen besteht. Für eine kontinuierliche Wiedergabe im ANLT wurden die Sätze des OLSA in randomisierter Reihenfolge mit Pausen zwischen 0,25 und 0,6 s dargeboten. Diese Werte liegen in einem Bereich, der von Campione und Veronis (2002) als mittlere Pausenlänge in fließend gesprochener Sprache bezeichnet wird. Als Hintergrundgeräusch wurde für alle Tests das Rauschen des OLSA gewählt. Es hat das gleiche Spektrum wie die Sprachaufnahmen des OLSA.

Für die technischen Messungen wurde folgendes Signal verwendet: Als Sprachsignal wurden 27 willkürlich gewählte Sätze kontinuierlich dargeboten, die zusammen einer Länge von 1,02 min entsprachen. Das OLSA-Rauschen wurde als Hintergrundgeräusch eingesetzt, das 7 s vor der Sprachwiedergabe begann. Durch den früheren Beginn konnte das Rauschen durch die Störgeräuschreduktion geschätzt werden. Zur Bestimmung des SNR wurden nur die Sprach- und Rauschsignale analysiert, die innerhalb der 1,02 min langen kontinuierlichen Sprachwiedergabe verarbeitet wurden. In Abb. 5.2 ist die zeitliche Abfolge der verwendeten Signale dargestellt.

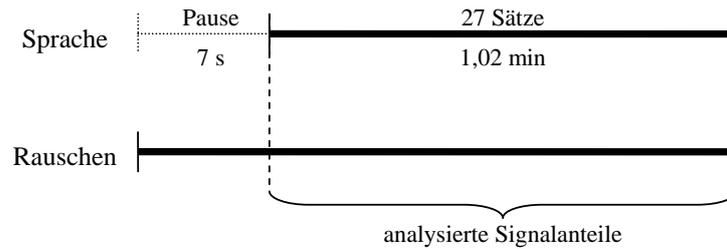


Abb. 5.2: Zur Analyse der tatsächlichen SNR-Verbesserung genutzte Signale.

5.3.3 Einkanalige Störgeräuschreduktionen

Mit den genannten Testverfahren wurden drei unterschiedliche einkanalige Störgeräuschreduktionen geprüft. Untersucht wurde ein optimaler Algorithmus, der über ein a priori Wissen des Hintergrundgeräusches verfügt, d. h., dass das Störgeräuschsignal getrennt von der Sprache vorliegt und keine Störgeräuschschätzung benötigt wird. Da Sprachsignal und Hintergrundgeräusch in realen Umgebungen am Eingang der Störgeräuschreduktion ein Gesamtsignal bilden, das nicht mehr getrennt werden kann, ist diese Verarbeitung unrealistisch. Das a priori Wissen wurde in einer Verstärkungsfunktion nach Wiener angewendet (Vary *et al.*, 1998). Dieses Filter ist danach optimal im Sinne Wiensers. Seine maximale theoretische Reduktion wurde auf 6 dB begrenzt. D. h., die Signale konnten durch die Verwendung eines Spectral Floor (Kroschel, 2004) in der Störgeräuschreduktion um maximal 6 dB abgesenkt werden. Im Folgenden wird dieser Algorithmus mit *Optimal* bezeichnet.

Außerdem wurden zwei realistische Algorithmen verwendet. Sie schätzen aus dem Gesamtsignal das Rauschen mit Hilfe des Minima Controlled Recursive Averaging (Cohen und Berdugo, 2002) und filtern das Signal nach der Methode der Spektralen Subtraktion der Leistungsdichtespektren (Vary *et al.*, 1998). Ihre maximale theoretische Reduktion wurde auf 6 dB und 8 dB begrenzt. Sie werden im Folgenden mit *Real6dB* und *Real8dB* bezeichnet. Zusätzlich zu den Messungen mit den Störgeräuschreduktionen *Optimal*, *Real6dB* und *Real8dB* wurden Messungen ohne Störgeräuschreduktion durchgeführt. Die Ergebnisse sind mit *NoAlgo* gekennzeichnet. Alle Störgeräuschreduktionen arbeiten mit der Weighted Overlap-Add-Methode, in der eine Fast-Fourier-Transformationslänge von 128 Samples mit 50%-iger Überlappung verwendet wird. Die Signale wurden mit einer Frequenz von 32000 Hz abgetastet.

Die Störgeräuschreduktionen verarbeiteten neben dem Gesamtsignal aus Sprache und Hintergrundgeräusch auch zeitgleich die separaten Sprach- und Rauschsignale mit den für das Gesamtsignal ermittelten Parametern. Dies kann zur exakten Bestimmung des SNR am Ausgang des Systems verwendet werden, da die Filterung linear geschieht. Aufgrund der getrennten Verarbeitung des Sprach- und Rauschsignals wurde die SNR-Verbesserung für jede Störgeräuschreduktion in Abhängigkeit vom SNR am Eingang ermittelt. Sie ist definiert als Differenz zwischen dem SNR in dB der Signale am Aus- und Eingang der Störgeräuschreduktion und wird im Folgenden tatsächliche SNR-

Verbesserung $\text{SNR}_{\text{OUT-IN}}$ genannt. Für diese Berechnung wurde das in Kap. 5.3.2 gezeigte Signal verwendet. In Abb. 5.3 ist die tatsächliche SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ dargestellt. Sie erreicht ihr Maximum, wenn der Eingangs-SNR SNR_{IN} zwischen 2 und 6 dB SNR liegt. Die Maxima liegen bei 3,3 dB SNR (*Optimal*), 1,9 dB SNR (*Real8dB*) und 1,6 dB SNR (*Real6dB*). Die tatsächliche SNR-Verbesserung ist bei allen Störgeräuschreduktionen niedriger als die theoretisch mögliche Reduktion, die nur die maximale Pegelverringerung beschreibt und nicht die tatsächliche SNR-Verbesserung berücksichtigt.

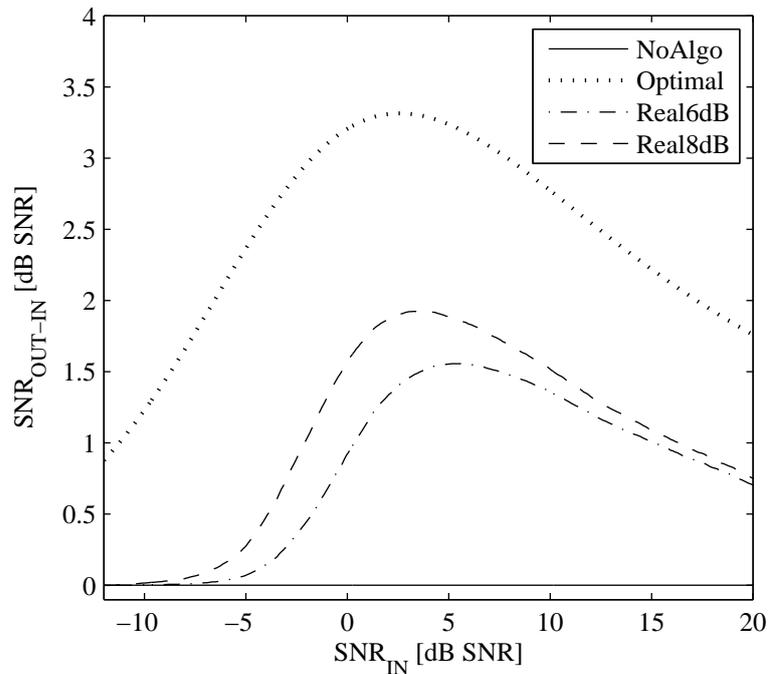


Abb. 5.3: Tatsächliche SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ der Störgeräuschreduktion, aufgetragen über dem SNR am Eingang SNR_{IN} .

Da einkanalige Störgeräuschreduktionen auf das gestörte Sprachsignal eine Dämpfung pro Frequenzband anwenden, wird der Pegel der Sprache und des Rauschens verändert, wenn beide Signalanteile sich spektral überlappen. Damit der Pegel der Sprache im ANLT konstant bleibt, ist es notwendig, diesen Pegelverlust auszugleichen. Dazu wurde das getrennt verarbeitete Sprachsignal genutzt. Die Pegel des Sprachsignals vor und nach der Störgeräuschreduktion wurden verglichen und daraus die benötigte Verstärkung berechnet. Um eine sprunghafte Veränderung des Gesamtsignalpegels zu verhindern, wurde die berechnete Verstärkung mit einem Fenster der Länge 1 s geglättet und auf das Ausgangssignal der Störgeräuschreduktion angewendet.

5.3.4 Perzeptives Messverfahren: ANLT

Vor Beginn der Untersuchungen wurde mit jeder Versuchsperson eine Tonaudiometrie durchgeführt. Mit Versuchspersonen, denen das Sprachmaterial nicht in ausreichendem Maße bekannt war, wurde vor Beginn der Messungen geübt. Dazu wurden zufällig drei Testlisten mit 30 Sätzen ausgewählt und bei einem konstanten Pegel ohne Hintergrundgeräusch präsentiert. Gewählt wurde ein Pegel, bei dem die Versuchspersonen die Sätze

ohne Anstrengung verstehen konnten. Sie hatten dann die Aufgabe, die Sätze zu wiederholen. Fehler wurden korrigiert, indem den Versuchspersonen die richtige Antwort vorgesprochen wurde.

Basierend auf den Beschreibungen nach Nabelek *et al.* (1991, 2006) wurde mit jeder Versuchsperson zuerst der MCL für Sprache bestimmt. Dazu wurde sie angewiesen, den Pegel der Sprache lauter als angenehm, danach leiser als angenehm und abschließend auf einen angenehmen Pegel (MCL) einzustellen, dem sie über eine längere Zeit zuhören würde. Der Pegel des ersten Satzes betrug 30 dB(A). Zur Einstellung des lauten und leisen Pegels wurden 5 dB-Schritte und zur Einstellung des MCL wurden 2 dB-Schritte verwendet. Diese Einstellungen wurden dreimal vorgenommen und der Median der MCL-Werte ermittelt. Die Instruktion zur Bestimmung des MCL ist in Anhang B.3 dargestellt.

Zur Bestimmung des BNL wurde die Sprache konstant beim Median der MCL-Werte und zusätzlich das Hintergrundgeräusch bei einem Pegel von 30 dB(A) präsentiert. Der Pegel des Hintergrundgeräusches wurde dann durch die Versuchsperson verändert. Zuerst wählte sie eine Lautstärke, bei der sie die Sätze nicht mehr verstehen konnte, danach sollte sie den Pegel verringern, bis die Sätze deutlich zu verstehen waren. Zuletzt wählte sie die maximale Lautstärke des Hintergrundgeräusches, die sie über einen längeren Zeitraum tolerieren würde, ohne sich belästigt zu fühlen. Sie sollte den Sätzen zuhören können, ohne zu ermüden. Dabei sollte sie sich vorstellen, im Radio einer Dokumentation zu folgen und auf den Pegel eines Hintergrundgeräusches (z. B. eines Staubsaugers) direkt Einfluss nehmen zu können. Die Instruktion zur Bestimmung des BNL und ANL ist in Anhang B.4 dargestellt. Auch hier wurde der maximale und minimale Pegel des Rauschens mit einer Schrittgröße von 5 dB und der tolerierbare Pegel mit einer Schrittgröße von 2 dB eingestellt. Aus der Differenz zwischen MCL und BNL wurde der ANL berechnet ($ANL = MCL - BNL$). Für jede der vier Testsituationen (*Optimal*, *Real6dB*, *Real8dB* und *NoAlgo*) wurde der BNL dreimal ermittelt. Die verwendeten Testsituationen wurden innerhalb der Messung randomisiert. Insgesamt führte die Versuchsperson die Bestimmung des ANL also zwölfmal durch. Für jede Situation wurde der Median der ANL-Werte gebildet.

Zur Einstellung des Pegels im ANLT wurde den Versuchspersonen eine graphische Oberfläche auf einem Bildschirm dargeboten (s. Anhang B.4). Dort wurde die aktuelle Instruktion (z. B. „...lauter als angenehm...“ oder „...leiser als angenehm...“) präsentiert und die Versuchspersonen hatten die Möglichkeit, mit Hilfe von Schaltflächen die Lautstärke der Signale zu verändern. Die Schaltflächen konnten mit der Maus des PCs ausgewählt werden. Hatte die Versuchsperson eine PegelEinstellung beendet, so bestätigte sie die Einstellung mit einem OK-Button und die Instruktion für den nächsten Messschritt im ANLT wurde präsentiert. Der OK-Button wurde erst 2 s nach der letzten Pegelveränderung auswählbar, damit das Einschwingen der Störgeräuschreduktion abgeschlossen war, bevor die Versuchsperson antwortete.

5.3.5 Technische Messverfahren

Neben dem bisher genannten perzeptiven Messverfahren wurden auch signalbasierte Untersuchungen durchgeführt. Zur Untersuchung der Störgeräuschreduktionen wurden die technischen Maße ΔSNR und MTF berechnet.

SNR nach Hagerman und Olofsson

Für die Ermittlung des ΔSNR nach Hagerman und Olofsson (2004) wurden zwei unterschiedliche Messungen mit Sprache und Störgeräusch durchgeführt. In beiden Messungen wurden die Signale bei definierten SNR-Werten zwischen -12 und 22 dB SNR mit einer Schrittgröße von 1 dB den verschiedenen Störgeräuschreduktionen präsentiert. In einer der beiden Messungen wurde das Rauschen invertiert. Die Ausgangssignale aus der Störgeräuschreduktion dieser beiden Messungen wurden sowohl addiert als auch subtrahiert, wodurch eine Trennung von Sprache und Hintergrundgeräusch möglich war. Aus den Übertragungsfunktionen für Sprache und Hintergrundgeräusch wurde die Veränderung des SNR berechnet. Um einen skalaren Wert zu erhalten, wurden die Werte mit dem Speech Intelligibility Index (SII) durch Anwendung der Tabelle 1 der Norm ANSI S 3.5 - 1997 (1997) gewichtet. Ergebnis war der ΔSNR . Die Verarbeitung der Signale wurde ohne Pegelausgleich, wie in Kap. 5.3.3 beschrieben, vorgenommen. Darauf wurde verzichtet, da nur Signale am Ausgang der Störgeräuschreduktion miteinander verglichen wurden und die Berechnung des ΔSNR eine Normierung des Pegels beinhaltet.

Modulationstransferfunktion

Zur Ermittlung der MTF wurde die Division zwischen dem Kreuzleistungsdichtespektrum des Ein- und Ausgangssignals der Störgeräuschreduktion und dem Autoleistungsdichtespektrum des Eingangssignals berechnet und normiert (Drullmann *et al.*, 1994). Die Signale wurden deshalb bei definierten SNR-Werten zwischen -12 und 22 dB SNR mit einer Schrittgröße von 1 dB präsentiert. Nur der Realteil der MTF wurde im weiteren Verfahren betrachtet. Diese Vereinfachung war unter der Annahme möglich, dass nur phasensynchrone Modulationen berücksichtigt werden sollten. Auch diese Werte wurden mit Hilfe des SII gewichtet, um einen skalaren MTF-Wert zu erhalten. Der Pegelverlust der Signale durch die Störgeräuschreduktion wurde ebenfalls nicht mit dem in Kap. 5.3.3 beschriebenen Verfahren ausgeglichen. Dieses war möglich, da in der Berechnung der MTF nur die Veränderungen der Modulationstiefe betrachtet werden, die unabhängig vom Gesamtpegel der Signale sind und die Berechnung der MTF eine Normierung des Pegels beinhaltet.

5.4 Ergebnisse

Perzeptives Maß: ANL

Die Einzelergebnisse des ANLT für jede Situation und jede Versuchsperson sind in Anhang E.1 aufgelistet und in Abb. 5.4a für die Normalhörenden und Abb. 5.4b für die Schwerhörigen dargestellt. Die Ergebnisse zeigen, dass durch die Verwendung der optimalen Störgeräuschreduktion in der Regel eine Verringerung des ANL erreicht wurde, d. h., die Versuchspersonen tolerierten einen höheren Pegel des Hintergrundgeräusches mit *Optimal* als ohne Störgeräuschreduktion (*NoAlgo*). Werden die realen Störgeräuschreduktionen (*Real6dB* und *Real8dB*) verwendet, kann kein eindeutiger Trend erkannt werden. Einige Versuchspersonen stellten einen höheren Pegel des Hintergrundgeräusches ein, andere dagegen nicht. Unterschiede zur Situation ohne Störgeräuschreduktion (*NoAlgo*) wurden mit einem Wilcoxon Vorzeichen-Rangtest auf Signifikanz bei einem Niveau von $\alpha = 0,05$ geprüft und für die Normalhörenden mit *Optimal* ($p = 0,004$) und *Real8dB* ($p = 0,023$) ermittelt. Die Ergebnisse der Schwerhörigen zeigen nur für die optimale Störgeräuschreduktion einen signifikanten Unterschied ($p = 0,001$). In Anhang E.2 sind die ermittelten Irrtumswahrscheinlichkeiten aller Vergleiche aufgelistet.

Damit der Nutzen durch eine Störgeräuschreduktion unabhängig von den interindividuellen Streuungen betrachtet werden kann, wurde die individuelle Differenz zwischen dem medianen ANL in der Situation ohne Störgeräuschreduktion und den medianen ANL-Werten mit Störgeräuschreduktion gebildet. Dieses Maß wird im Folgenden ΔANL genannt ($\Delta\text{ANL} = \text{ANL}_{\text{NoAlgo}} - \text{ANL}_{\text{Algo}}$). Die Ergebnisse sind für die normalhörenden Versuchspersonen in Abb. 5.5a und für die schwerhörigen Versuchspersonen in Abb. 5.5b dargestellt. Ein größerer Nutzen der Störgeräuschreduktion führt zu einem höheren ΔANL . Die Normalhörenden erreichten mit der optimalen Störgeräuschreduktion einen medianen ΔANL von 3 dB SNR, Schwerhörige von 4 dB SNR. Mit realen Störgeräuschreduktionen erreichen die Normalhörenden einen ΔANL von 1 dB SNR (*Real6dB*) bzw. 2 dB SNR (*Real8dB*). Schwerhörige wählten ΔANL -Werte von 0 dB SNR (*Real6dB*) bzw. 1 dB SNR (*Real8dB*). Die maximale tatsächliche SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ von 3,3 dB SNR (*Optimal*), 1,9 dB SNR (*Real8dB*) und 1,6 dB SNR (*Real6dB*, s. Abb. 5.3) konnte somit annähernd erreicht werden.

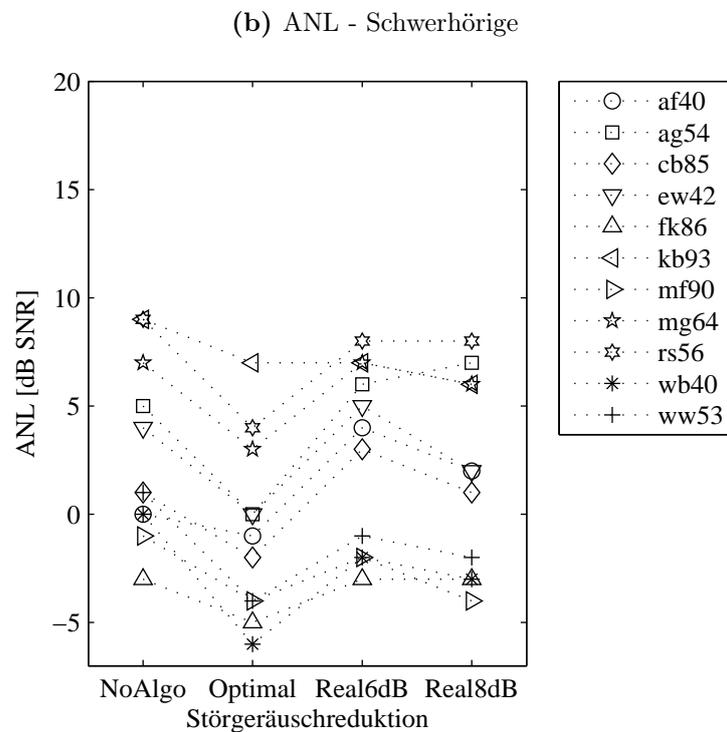
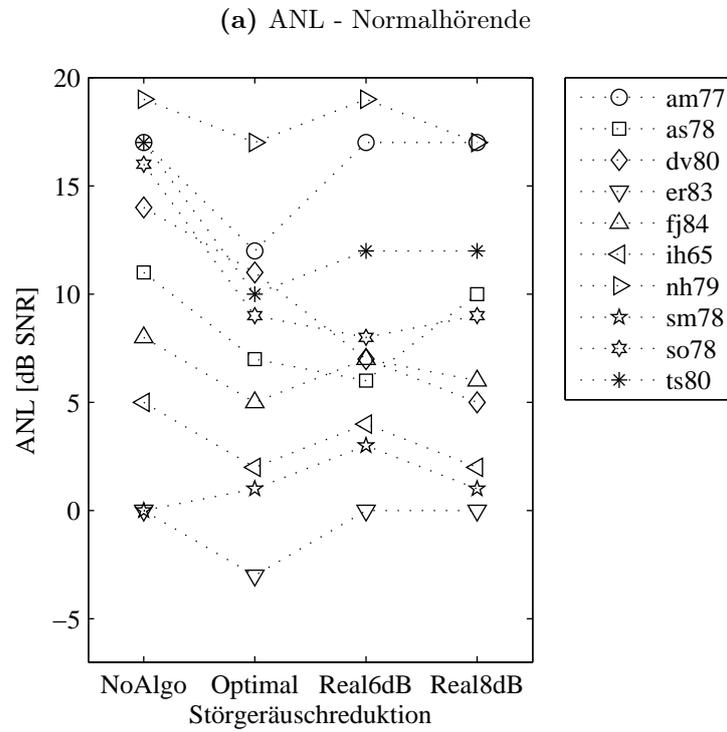


Abb. 5.4: ANL-Werte der (a) Normalhörenden und der (b) Schwerhörigen.

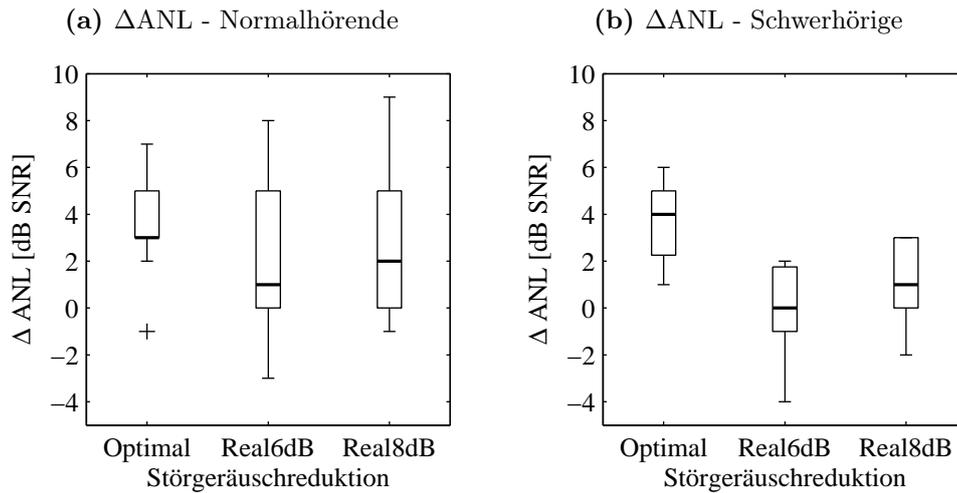


Abb. 5.5: Nutzen der unterschiedlichen Störgeräuschreduktionen, gemessen im ANLT für (a) Normalheerende und (b) Schwerhörige.

Um den Zusammenhang zwischen den ΔANL -Werten und der tatsächlichen SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ genauer zu untersuchen, wurde zu jedem eingestellten ΔANL -Wert der $\text{SNR}_{\text{OUT-IN}}$ -Wert bei gleichem SNR am Eingang der Störgeräuschreduktion ermittelt. Für den ΔANL wurden die ANL-Werte, gemessen mit der jeweiligen Störgeräuschreduktion, als SNR-Werte am Eingang gewählt. Die gewählten $\text{SNR}_{\text{OUT-IN}}$ -Werte wurden mit einem Spearmanschen Korrelationskoeffizienten r_{sp} mit den ΔANL -Werten verglichen. Seine Irrtumswahrscheinlichkeit wurde mit dem Permutationstest ermittelt. Der Test prüft die Hypothese, dass die Korrelation signifikant unterschiedlich von $r_{\text{sp}} = 0$ ist. Keine der Korrelationen ist signifikant bei einem Signifikanzniveau von $\alpha = 0,05$. Anhang E.4 zeigt eine Gegenüberstellung der Ergebnisse des ANLT mit der tatsächlichen SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ und die ermittelten Korrelationen und Irrtumswahrscheinlichkeiten.

Technische Maße

In Abb. 5.6a und Abb. 5.6b sind der ΔSNR nach Hagerman und Olofsson (2004) und die MTF nach Drullmann *et al.* (1994) für die unterschiedlichen Störgeräuschreduktionen über dem SNR am Eingang der Störgeräuschreduktion SNR_{IN} aufgetragen. Die Maxima aller Kurven des ΔSNR liegen bei positiven SNR_{IN} . Die optimale Störgeräuschreduktion erreicht die höchsten ΔSNR -Werte. Die ΔSNR -Werte von *Real8dB* liegen über denjenigen von *Real6dB*. Die in Abb. 5.6b dargestellte MTF der Störgeräuschreduktionen ist größer eins. Damit haben alle Störgeräuschreduktionen die Modulationstiefe der Signale vergrößert. Die maximalen MTF-Werte der Störgeräuschreduktionen *Optimal* und *Real8dB* sind etwa gleich, liegen jedoch bei unterschiedlichen SNR_{IN} -Werten. Die Werte von *Real6dB* sind niedriger als die von *Real8dB* und übersteigen die Werte von *Optimal* bei hohen SNR_{IN} -Werten.

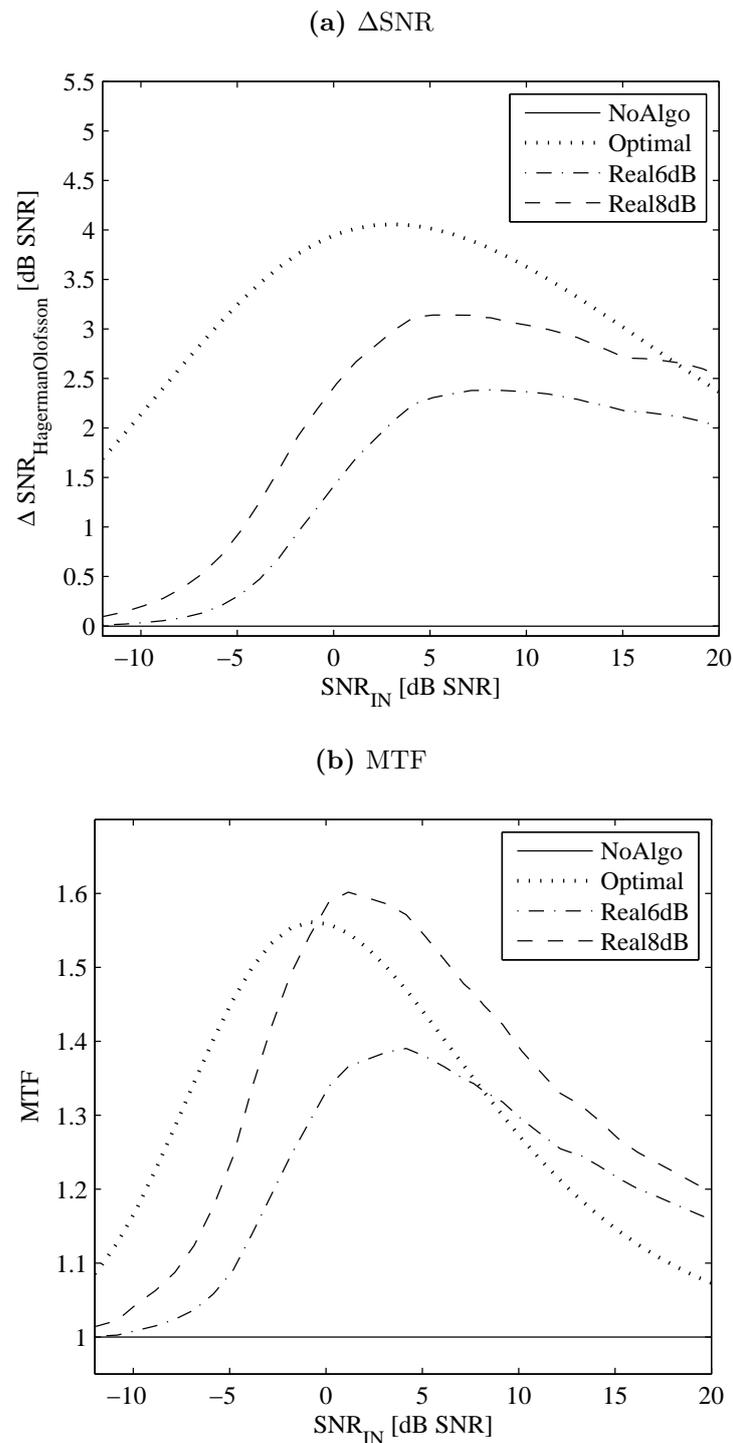


Abb. 5.6: Δ SNR und MTF der verwendeten Störgeräuschreduktionen, aufgetragen über dem SNR am Eingang der Störgeräuschreduktion SNR_{IN} .

Zum Vergleich der technischen Maße mit der tatsächlichen SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ wurden die Werte zusammen dargestellt. Abb. 5.7 zeigt in der linken Spalte die ermittelten Δ SNR-Werte und in der rechten Spalte die MTF-Werte zusammen mit $\text{SNR}_{\text{OUT-IN}}$ für jede Störgeräuschreduktion. Für diesen Vergleich wurden die MTF-Werte logarithmiert und auf den maximalen $\text{SNR}_{\text{OUT-IN}}$ -Wert normiert. Beide technischen Maße weichen von der tatsächlichen SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ ab, der Δ SNR vor allem bei hohen SNR-Werten am Eingang SNR_{IN} der realen Störgeräuschreduktionen.

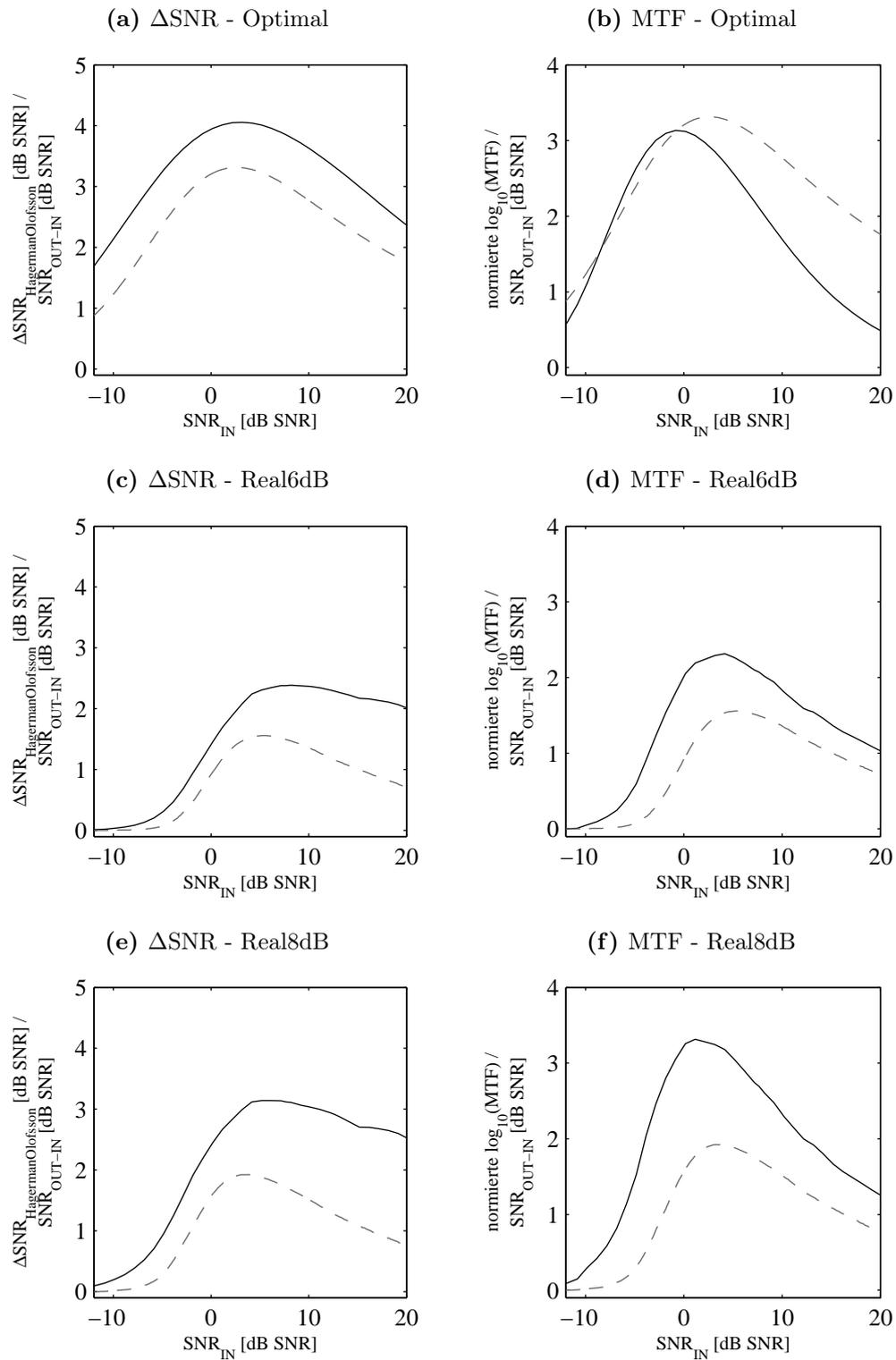


Abb. 5.7: ΔSNR und normierte und logarithmierte MTF aufgetragen über dem SNR am Eingang der Störgeräuschreduktion SNR_{IN} . Die durchgezogenen Linien in den Teilabbildungen zeigen die Ergebnisse der technischen Maße ΔSNR (linke Spalte) und MTF (rechte Spalte) für die Störgeräuschreduktionen *Optimal*, *Real6dB* und *Real8dB*. Die unterbrochene Linie stellt die tatsächliche SNR-Verbesserung SNR_{OUT-IN} jeder Störgeräuschreduktion aus Abb. 5.3 dar.

Um die Δ SNR- und die MTF-Werte mit den $\text{SNR}_{\text{OUT-IN}}$ -Werten zu vergleichen, wurde der Spearmansche Korrelationskoeffizient r_{sp} und die Irrtumswahrscheinlichkeit p berechnet und in Tab. 5.1 aufgeführt. Die Korrelationen von Δ SNR liegen zwischen 0,9 und 0,99 und die der MTF zwischen 0,83 und 0,94. Alle Koeffizienten sind signifikant auf einem Niveau von $\alpha = 0,01$.

Tab. 5.1: Spearmanscher Korrelationskoeffizient r_{sp} und Irrtumswahrscheinlichkeit p für den Zusammenhang zwischen Δ SNR und der tatsächlichen SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ durch die Störgeräuschreduktion.

| Störgeräuschreduktion | Δ SNR r_{sp} | MTF r_{sp} |
|-----------------------|---------------------------------|------------------------|
| Optimal | 0,996 | 0,832 |
| Real6dB | 0,917 | 0,944 |
| Real8dB | 0,904 | 0,926 |

Vergleich des perzeptiven und der technischen Maße

Auch zum Vergleich der Δ ANL- mit den Δ SNR- und MTF-Werten wurde der Spearmansche Korrelationskoeffizient r_{sp} berechnet. Dazu wurden linear interpolierte Werte des Δ SNR oder der normierten logarithmierten MTF verwendet, wenn keine Werte beim gleichen SNR_{IN} wie für Δ ANL zur Verfügung standen. Die Korrelationskoeffizienten sind in Tab. 5.2 zusammen mit der Irrtumswahrscheinlichkeit p eines Permutationstestes dargestellt. Die Korrelationen der Δ ANL-Werte der Normalhörenden und Δ SNR bei Verwendung der Störgeräuschreduktion *Real6dB* und *Real8dB* sind signifikant auf einem Niveau von $\alpha = 0,05$ und betragen 0,89 und 0,71. Außerdem korrelieren die Werte der Schwerhörigen beim Vergleich Δ ANL und MTF signifikant. Die Korrelation beträgt -0,72. Alle anderen Korrelationen der Normal- und Schwerhörigen sind nicht signifikant. Anhang F zeigt eine grafische Gegenüberstellung der Werte.

Tab. 5.2: Spearmanscher Korrelationskoeffizient r_{sp} und Irrtumswahrscheinlichkeit p für den Zusammenhang zwischen Δ ANL und Δ SNR nach der Störgeräuschreduktion.

| Untersucher Vergleich | Störgeräuschreduktion | Normalhörende | | Schwerhörige | |
|-----------------------------|-----------------------|-----------------|-------|-----------------|-------|
| | | r_{sp} | p | r_{sp} | p |
| Δ ANL - Δ SNR | Optimal | -0,075 | 0,836 | 0,060 | 0,860 |
| | Real6dB | 0,889 | 0,001 | -0,079 | 0,818 |
| | Real8dB | 0,709 | 0,022 | -0,267 | 0,428 |
| Δ ANL - MTF | Optimal | -0,320 | 0,367 | -0,130 | 0,703 |
| | Real6dB | 0,115 | 0,753 | -0,724 | 0,012 |
| | Real8dB | -0,116 | 0,749 | -0,413 | 0,207 |

5.5 Diskussion

Ziel dieser Untersuchung war, das Verfahren des ANLT als Untersuchungsform für Störgeräuschreduktionen zu prüfen. Die ANL-Ergebnisse zeigen, dass mit den Störgeräuschreduktionen keine SNR-Verbesserung erreicht werden konnte, die der maximalen theoretischen Reduktion des Hintergrundgeräusches entsprach. Dies wurde auch nicht erwartet, da die tatsächliche SNR-Verbesserung durch die Störgeräuschreduktionen geringer ist. Sie konnte z. T. mit den ANL-Ergebnissen nachgebildet werden. Der größte Nutzen wurde durch den optimalen Algorithmus erreicht. Dies legt nahe, dass der ANLT grundsätzlich als Methode zur Untersuchung von Störgeräuschreduktionen geeignet ist. Die realen Algorithmen mit einer maximalen Reduktion von 6 oder 8 dB zeigen eine geringe Verbesserung bzw. eine Tendenz, dass sich die Ergebnisse verbessern. Dabei sind die Ergebnisse der Schwerhörigen geringgradig schlechter als die der Normalhörenden. Damit konnten die ANL-Ergebnisse von Mueller *et al.* (2006) nicht in ihrer Deutlichkeit bestätigt werden. Ursächlich für diese Ergebnisse könnten die gemessenen ANL-Werte der Versuchspersonen sein, die eine große Streuung und keine signifikanten Korrelationen mit der tatsächlichen SNR-Verbesserung aufweisen. Diese Abweichungen könnten durch die verwendete Schrittgröße hervorgerufen sein, die eine Einstellung der z. T. geringen Unterschiede zwischen den Störgeräuschreduktionen nicht möglich macht. Außerdem ist das im ANLT gewählte Kriterium des „tolerierbaren“ Pegels vermutlich keine eindeutige Schwelle. Es ist möglich, dass dieses Kriterium nicht klar durch die Versuchspersonen definiert werden kann.

Außerdem wurden in dieser Studie die technischen Maße Δ SNR nach Hagerman und Olofsson (2004) und MTF nach Drullmann *et al.* (1994) verwendet, um den Nutzen durch Störgeräuschreduktionen zu beschreiben. Die Korrelationen zwischen den Δ SNR-Werten und der tatsächlichen SNR-Verbesserung sind signifikant für die unterschiedlichen Störgeräuschreduktionen und liegen bei 0,9-0,99. Vor allem bei hohen SNR-Werten am Eingang der realen Störgeräuschreduktionen weichen die tatsächliche SNR-Verbesserung und der Δ SNR aber deutlich voneinander ab. Ursache für diese deutlichen Unterschiede könnte der hohe SNR sein. Sprache und Rauschen unterscheiden sich bei hohen SNR-Werten deutlich in ihren Pegeln, die Veränderung des Rauschpegels durch die Störgeräuschreduktion ist relativ dazu nur gering. Einschwingvorgänge der Störgeräuschreduktionen oder andere Artefakte sind nicht linear und werden vom Δ SNR nicht berücksichtigt.

Die Ergebnisse der MTF, die außerdem zur Beschreibung des Nutzens der Störgeräuschreduktion genutzt wurde, sind größer eins. Für alle Störgeräuschreduktionen wurde also eine Erhöhung der Modulationstiefe ermittelt. Die MTF überschätzte die Störgeräuschreduktion *Real8dB* deutlich. Diese Störgeräuschreduktion verarbeitet die Signale mit Artefakten, die vor allem bei niedrigen SNR-Werten am Eingang deutlich zu hören sind. Diese werden vermutlich durch die MTF als Modulationsveränderungen interpretiert, was zu hohen Werten führt. Dennoch sind signifikante Korrelationen von 0,84-0,94 zwischen der tatsächlichen SNR-Verbesserung und der normierten und logarithmierten MTF vorhanden. Der Δ SNR nach Hagerman und Olofsson (2004) scheint aufgrund sei-

ner höheren Korrelation jedoch eher geeignet zu sein, die tatsächliche SNR-Verbesserung durch eine Störgeräuschreduktion zu beschreiben.

Bei allen Betrachtungen ist zu berücksichtigen, dass für die Ermittlung eines skalaren Wertes der technischen Maße eine Gewichtung mit dem SII vorgenommen wurde. Diese Gewichtung wird für den ΔSNR vorgeschlagen und wurde in dieser Studie für die MTF übernommen. Eine Gewichtung der Werte bei der Bestimmung der tatsächlichen SNR-Verbesserung wurde nicht berücksichtigt. Da das Rauschsignal jedoch das gleiche Spektrum wie die Sprache besitzt, sind die Auswirkungen vermutlich gering. Dies sollte in weiteren Studien jedoch geprüft werden.

Ein Vergleich des ΔSNR und der MTF mit dem perzeptiven Maß ΔANL zeigt wenige signifikante Korrelationen. Für einen Vergleich zwischen dem ΔANL und dem ΔSNR für Normalhörende bei realen Störgeräuschreduktionen wurden Korrelationen von 0,89 und 0,71 ermittelt. Die MTF korrelierte mit den ΔANL -Werten der Schwerhörigen bei Verwendung der Störgeräuschreduktion *Real6dB* ($r_{\text{sp}} = -0,72$). Ein Vergleich der technischen Maße mit dem perzeptiven Maß ist aber dennoch schwierig, da nicht für alle Messungen Korrelationen ermittelt werden konnten. Die starke interindividuelle Streuung der ANL-Werte zeigt, dass dieses Maß sehr von den Versuchspersonen abhängig ist. Dies ist beim Vergleich mit den technischen Maßen zu berücksichtigen und eine Verallgemeinerung der Ergebnisse fällt daher schwer.

Eine optimale technische Methode zur Untersuchung des Nutzens von Störgeräuschreduktionen sollte die Artefakte der Störgeräuschreduktionen beachten. Diese könnte evtl. durch eine technische Methode, die die menschliche Perzeption berücksichtigt, modelliert werden. Dazu könnte z. B. das Oldenburger Perzeptionsmodell zur Qualitätsvorhersage von Signalen (Huber, 2003) verwendet werden. In dieser Methode wird zuerst ein auditives Verarbeitungsmodell auf ein Referenz- und Testsignal angewendet. Die Ausgänge des Modells werden durch Korrelationskoeffizienten verglichen und beschreiben im Qualitätsmaß Perceptual Similarity Measure die Ähnlichkeit der Signale.

5.6 Schlussfolgerungen

Unter der Berücksichtigung, dass technische Maße mit dem SII gewichtet wurden und diese Gewichtung nicht bei der tatsächlichen SNR-Verbesserung oder dem ANL beachtet wurde, lässt diese Studie folgende Schlüsse zu:

1. Zur Untersuchung von einkanaligen Störgeräuschreduktionen mit Sprache im Störgeräusch kann der ANLT generell verwendet werden. Probleme bereitet die große Streuung der ANL-Werte. Verschiedene Versuchspersonen verwenden verschiedene ANL-Werte und damit verschiedene SNR-Werte am Eingang der Störgeräuschreduktion. Die Störgeräuschreduktion erreicht bei unterschiedlichen SNR-Werten am Eingang unterschiedliche SNR-Verbesserungen. Außerdem sind die geringen Unterschiede der Störgeräuschreduktionen problematisch. Sie sind z. T. geringer als die verwendete Schrittgröße.

2. ΔSNR nach Hagerman und Olofsson (2004) und MTF nach Drullmann *et al.* (1994) eignen sich nur bedingt, die tatsächliche SNR-Verbesserung durch die Störgeräuschreduktion zu beschreiben.
3. Ein Vergleich der technischen Maße ΔSNR und MTF mit dem perceptiven Maß ANL ist schwierig. Die technischen Maße korrelieren nur vereinzelt mit den Ergebnissen des ANLT.

Kapitel 6

Zusammenfassung und Ausblick

Ziel dieser Arbeit war es, die Methode des Acceptable Noise Level Tests zu analysieren und seine Eignung zur Nutzendokumentation von Störgeräuschreduktionen zu untersuchen. Störgeräuschreduktionen sollen das Signal-Rausch-Verhältnis (SNR) zwischen Nutz- und Störsignal und damit Sprachverständlichkeit und Klangqualität verbessern. Eingangssignal einer einkanaligen Störgeräuschreduktion ist das gestörte Nutzsinal, das mit Hilfe mehrerer Verarbeitungsschritte entstört wird. Dazu ist zuerst die Leistung des Störsignals zu schätzen. Verwendet werden können dazu z. B. die Methoden Minimum Statistics (Martin, 1993, 1994, 2001, 2003), Amplitudenmodulationsspektrogramm (Tchorz und Kollmeier, 2003), Voice Activity Detection und Periodogramm (Appiah *et al.*, 2005) oder Minima Controlled Recursive Averaging (Cohen und Berdugo, 2002). Die geschätzte Rauschleistung wird dann in einer Verstärkungsregel, wie z. B. der Spektralen Subtraktion, dem Optimalfilter (Vary *et al.*, 1998) oder der Störgeräuschreduktion nach Ephraim und Malah (1984, 1985) verwendet. Der Nutzen der Störgeräuschreduktionen kann mit unterschiedlichen perzeptiven und technischen Tests untersucht werden, die Sprache im Störgeräusch als Signale verwenden.

Eine perzeptive Methodik zur Untersuchung von Sprache im Störgeräusch ist der Acceptable Noise Level Test (ANLT, Nabelek *et al.* (1991, 2006)). Er ermittelt einen akzeptierbaren SNR für Sprache im Hintergrundgeräusch, den Acceptable Noise Level (ANL). Dieser ist definiert als Differenz zwischen dem angenehmen Pegel der Sprache (engl. *Most Comfortable Level*, *MCL*) und dem maximalen Pegel des Hintergrundgeräusches, den Versuchspersonen über längere Zeit beim Hören von Sprache akzeptieren würden (engl. *Background Noise Level*, *BNL*). Personen mit einem niedrigen ANL akzeptieren einen lauterer Pegel des Hintergrundgeräusches, während sie der Sprache zuhören, als Personen mit einem hohen ANL, die wenig Hintergrundgeräusch tolerieren. Der ANLT kann mit der häufig verwendeten Schrittgröße 5 bzw. 2 dB, einem Startpegel von 30 dB(A) und der aufsteigenden Instruktion verwendet werden. Seine signifikante Test-Retest-Reliabilität ist mit 0,87 hoch. Sie kann erhöht werden, indem bei Messwiederholungen zuerst der MCL bestimmt wird und darauf die BNL-Bestimmungen bei gleichem MCL durchgeführt werden. Generell sind drei Messwiederholungen zur Berücksichtigung von Lerneffekten zu empfehlen. Außerdem zeigen die Ergebnisse dieser Arbeit, dass ANL-Werte für Stimmengewirr von zwölf Sprechern höher liegen als mit

anderen Hintergrundgeräuschen (z. B. sprachsimulierendes Rauschen, Aufnahmen einer Maschine oder ein Cafeteria-Geräusch). Dies Ergebnis ist nicht vergleichbar mit den Ergebnissen der Literatur (Nabelek *et al.*, 1991; Freyaldenhoven *et al.*, 2006c), in denen im Vergleich mit einem sprachsimulierenden Rauschen niedrigere ANL-Werte erreicht wurden. Außerdem wurde ermittelt, dass eine Übung des Sprachmaterials bei konstantem Darbietungspegel vor Beginn des ANLT keinen Einfluss auf dem ANL hat. Ebenso hat die Durchführung einer SRT-Ermittlung mit dem OLSA unmittelbar vor dem ANLT keinen Einfluss auf den ANL.

Naheliegende perzeptive Untersuchungsformen der Sprachverbesserung durch Störgeräuschreduktionen sind Sprachverständlichkeitstests im Störgeräusch wie der Oldenburger Satztest (OLSA, Wagener *et al.* (1999a,b,c); Wagener (2003)) und der Just Follow Conversation Test (JFCT, Hawkins und Stevens (1950); Larsby und Arlinger (1994); Borg *et al.* (1998); Magnusson *et al.* (2001); Neumann und Jensen (2001); Latzel *et al.* (2003)). Mit dem OLSA wurden Sprachverständlichkeitsschwellen 50%-iger Satzverständlichkeit (engl. *Speech Reception Threshold, SRT*) bestimmt. Mit dem JFCT wurden subjektive Sprachverständlichkeitsschwellen (JFC-Wert) ermittelt, indem die Versuchspersonen das Störgeräusch regelten, bis sie die Sätze gerade eben noch verstehen konnten. Ergebnisse dieser Arbeit zeigen, dass der JFCT sich nur für Messungen mit kurzem zeitlichen Abstand eignet und mindestens neun Messwiederholungen durchgeführt werden sollten. Außerdem zeigen die Ergebnisse die Tendenz, dass der JFC-Wert durch eine SRT-Ermittlung mit dem OLSA beeinflusst wird. Ein Wert, der die wahrgenommene Sprachverständlichkeit, gemessen mit dem JFCT, mit der tatsächlichen Sprachverständlichkeit, gemessen mit dem OLSA, vergleicht, ist die Perceptual Performance Discrepancy (PPDIS, Saunders und Cienkowski (2002)). Sie ist definiert als Differenz zwischen JFC-Wert und SRT des OLSA. Zur Untersuchung der Verbesserung der Signalqualität durch die Störgeräuschreduktionen wurden Paarvergleiche durchgeführt. Dafür wurden Signalpaare bei gleichem SNR dargeboten und die Versuchspersonen hatten den Paarling zu wählen, der die bessere Gesamtqualität aufwies.

Untersucht wurden mit dem OLSA, dem JFCT, der PPDIS, dem Paarvergleich und dem ANLT drei einkanalige Störgeräuschreduktionen. Ein Algorithmus verwendete ein a priori Wissen des Hintergrundgeräusches und eine Verstärkungsfunktion nach Wiener und seine maximale Reduktion wurde auf 6 dB SNR begrenzt. Zwei weitere Algorithmen schätzten das Rauschen mit Hilfe des Minima Controlled Recursive Averaging und filterten das Signal nach der Methode der Spektralen Subtraktion der Leistungsdichtespektren. Ihre maximale Reduktion wurde auf 6 bzw. 8 dB SNR beschränkt. Die Algorithmen wurden mit Situationen ohne Störgeräuschreduktion verglichen. Die Ergebnisse zeigten für normal- und schwerhörige Versuchspersonen, dass der ANLT zur Untersuchung von Störgeräuschreduktionen verwendet werden kann. Problematisch ist die interindividuelle Streuung des Acceptable Noise Level (ANL). Versuchspersonen, die unterschiedliche ANL-Werte einstellen, erreichen damit verschiedene SNR-Werte am Eingang der Störgeräuschreduktion. Diese reagiert mit unterschiedlichen SNR-Verbesserungen. OLSA und JFCT sind nicht geeignet, da im Test SNR-Werte erreicht werden, bei denen die Störgeräuschreduktion keine Verbesserung erzielt. Damit führt auch die PPDIS zu

keinem Ergebnis, das einen Nutzen durch Störgeräuschreduktionen beschreiben kann. Der Paarvergleich zeigt, dass die Schwerhörigen deutlich Situationen mit Störgeräuschreduktion bevorzugen. Diese Aussage trifft nicht auf die Ergebnisse der Normalhörenden zu, die Störgeräuschreduktionen mit Artefakten meiden. Eine Gegenüberstellung der Ergebnisse des ANLT mit dem Paarvergleich zeigt, dass die Akzeptanz des Hintergrundgeräusches bei den Schwerhörigen z. T. von ihrer Bevorzugung der Störgeräuschreduktion gegenüber Situationen ohne Störgeräuschreduktion abhängig ist. Die Ergebnisse zeigen auch niedrige signifikante Korrelationen zwischen den ermittelten Ergebnissen der Sprachverständlichkeitstests und dem ANLT. Der akzeptable Rauschpegel ist für die Schwerhörigen u. a. abhängig von der Sprachverständlichkeit.

Zur technischen Untersuchung von Störgeräuschreduktionen eignen sich Maße, die die SNR-Veränderung untersuchen. Dazu kann z. B. der Δ SNR nach Hagerman und Olofsson (2004) oder die Modulationstransferfunktion (MTF, Drullmann *et al.* (1994)) verwendet werden. Der Δ SNR ist ein Maß für die Verbesserung des SNR durch eine Störgeräuschreduktion und trennt durch Messungen mit Rauschen und invertiertem Rauschen die Signalanteile Sprache und Rauschen. Ihre SNR-Veränderung kann bestimmt werden. Die MTF untersucht dagegen die Veränderung der Modulationstiefe durch eine einkanalige Störgeräuschreduktion.

Geprüft wurden die o. g. Störgeräuschreduktionen mit Δ SNR und MTF und die Ergebnisse mit denen des ANLT verglichen. Beide technischen Maße eignen sich nur bedingt, die tatsächliche SNR-Verbesserung durch die Störgeräuschreduktion zu beschreiben. Sie berücksichtigen keine Artefakte der Störgeräuschreduktionen. Ein Vergleich der technischen Maße Δ SNR und MTF mit dem perzeptiven Maß ANL ist schwierig, da er stark streut. Die technischen Maße korrelieren nur vereinzelt mit den Ergebnissen des ANLT.

Eine optimale perzeptive Methode zur Untersuchung realer Störgeräuschreduktionen sollte die Messung bei einem konstanten und positiven SNR durchführen, bei dem eine maximale SNR-Verbesserung erreicht werden kann. Sie sollte ein einfaches Kriterium besitzen und ausreichend feine Schrittgrößen verwenden, mit denen die Unterschiede eingestellt werden können. Eine Lösung dieser Probleme ist evtl. die von Wittkop (1994, 2001) vorgeschlagene Methode. Er führte einen Vergleich zwischen einem durch Störgeräuschreduktionen verarbeiteten Referenzsignal und einem veränderlichen unverarbeiteten Signal durch. Die Versuchspersonen hatten die Aufgabe, den SNR des Referenzsignals mit dem unverarbeiteten Signal zu reproduzieren. Diese Methode ist zu prüfen.

Eine optimale technische Methode zur Untersuchung des Nutzens von Störgeräuschreduktionen sollte die Artefakte der Störgeräuschreduktionen beachten. Diese könnten evtl. durch eine technische Methode, die die menschliche Perzeption berücksichtigt, modelliert werden. Dazu könnte z. B. das Oldenburger Perzeptionsmodell zur Qualitätsvorhersage von Signalen (Huber, 2003) verwendet werden. In dieser Methode wird zuerst ein auditives Verarbeitungsmodell auf ein Referenz- und Testsignal angewendet. Die Ausgänge des Modells werden durch Korrelationskoeffizienten verglichen und beschreiben im Qualitätsmaß Perceptual Similarity Measure die Ähnlichkeit der Signale.

Anhang A

Messaufbau, Freifeldentzerrung und Kalibrierung

A.1 Messaufbau

Abb. A.1 zeigt schematisch den Messaufbau. Die Signale wurden mit Hilfe von Matlab erzeugt und über die Soundkarte (RME AD/DA-Interface ADI-2) an einen Abschwächer (Tucker Davis Technologies Attenuator PA 5) weitergegeben, dessen Abschwächung 0 dB betrug. Das Ausgangssignal wurde in einen Kopfhörerverstärker (Tucker Davis Technologies Headphone Driver HB7) eingespeist. Seine Verstärkung betrug ebenfalls 0 dB. An den Kopfhörerverstärker war ein circumauraler Kopfhörer (Sennheiser HDA 200) angeschlossen.

Zur Kalibrierung wurde der Kopfhörer auf einen Kuppler (GRAS Artificial Ear Type 43AA) aufgesetzt, dessen Mikrophon über einen Vorverstärker (GRAS Preamplifier Type 27a) mit einem Schallpegelmesser (Norsonic Type 118) verbunden war.

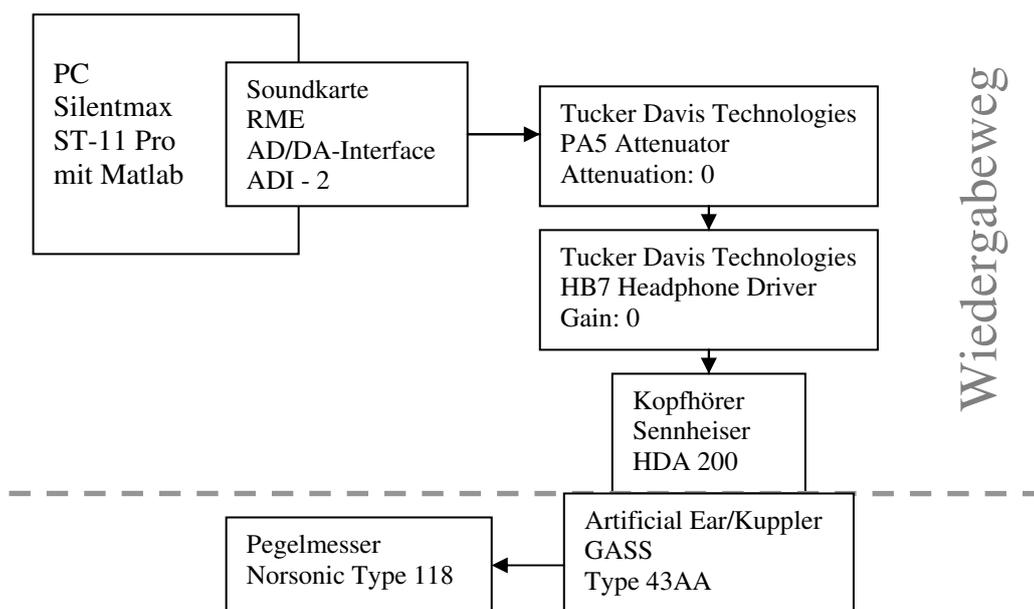


Abb. A.1: Messaufbau

A.2 Ermittlung der Freifeldentzerrung

Die Ermittlung der Freifeldentzerrung wurde nach DIN EN 606045-2 und DIN EN ISO 389-8 vorgenommen. Dazu wurde der Kopfhörer auf einen Kuppler (GRAS Artificial Ear Type 43AA) aufgesetzt. Das im Kuppler enthaltene Mikrofon wurde über einen Vorverstärker (GRAS Preamplifier Type 27a) mit einem Schallpegelmessers (Norsonic Type 118) verbunden und mit Hilfe eines Pistonphons kalibriert. Es wurden Terzrauschen bei den in DIN EN ISO 389-8 Anhang C beschriebenen Mittenfrequenzen erzeugt und über die Signalkette wiedergegeben. Der im Kuppler bei linearer Bewertung gemessene Pegel wurde dokumentiert und daraus das Kuppler-Übertragungsmaß G_C bestimmt. Dazu wurde die Differenz zwischen dem maximalen Pegel und allen anderen gemessenen Pegeln ermittelt, um die Abweichung der Pegel voneinander zu bestimmen. Aus den Werten der Tabelle C.1 der Norm EN ISO 389-8, die die Differenz zwischen Freifeldübertragungsmaß und Kuppler-Übertragungsmaß $G_F - G_C$ in dB enthält, kann dann durch Addition von G_C das Freifeldübertragungsmaß G_F ermittelt werden. Werden die Signale im Test mit einem Filter, das G_F wiedergibt, beaufschlagt, so sind sie freifeldentzerrt.

A.3 Kalibrierung

Um den Messaufbau zu kalibrieren, müssen für alle verwendeten Signale Kalibrierwerte ermittelt werden. Das OLSA-Rauschen eignet sich dabei auch zur Ermittlung der Kalibrierwerte der Sprache des OLSA, da der Frequenzgang und der Langzeitpegel des OLSA-Rauschens dem der Sprache gleich kommt. Das zu kalibrierende Signal wird mit Hilfe von Matlab mit einem $G_F - G_C$ (s. DIN EN ISO 328-9 Anhang C) beaufschlagt und über den Messaufbau wiedergegeben. Im Kuppler wird mit Hilfe des Pegelmessers der Pegel des Signals in dB(A) gemessen. Die Kalibrierkonstante K in dB ist beschrieben durch den internen RMS-Pegel des Signals RMS_{FS} und seinen erzeugten Pegel im Kuppler $RMS_{dB(A)}$ (s. A.1):

$$K = RMS_{dB(A)} - RMS_{FS} \quad (\text{A.1})$$

Die Kalibrierkonstante K kann dazu verwendet werden, den Pegel des Signals zu steuern. Die interne Verstärkung V_{int} des Signals, die notwendig ist, um einen gewünschten Pegel am Ausgang zu erzeugen, wird berechnet mit Hilfe des Pegels Ref , bei dem das Signal dargeboten werden soll, des internen RMS-Pegels des Signals RMS_{FS} und der Kalibrierkonstante K (s. A.2):

$$V_{int} = Ref - K - RMS_{FS} \quad (\text{A.2})$$

V_{int} wird in einen linearen Wert umgerechnet und mit dem Signal multipliziert. Am Ende der Darbietungskette kann so ein Signal bei dem Pegel Ref gemessen werden.

Anhang B

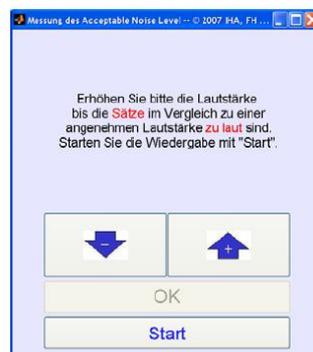
Instruktionen

B.1 Instruktion des ANLT

Annehmbarer Pegel des Hintergrundgeräusches

Ihnen werden gleich Sätze dargeboten, die von einem Sprecher vorgelesen werden. Als Erstes werden Sie gebeten, die Lautstärke der Sätze auf einen angenehmen Pegel einzustellen, d. h. etwa so, als ob Sie Radio hören würden. Dazu soll die Lautstärke erst lauter und danach leiser eingestellt werden. Danach wählen Sie bitte eine Lautstärke die Ihnen sehr angenehm ist.

Im Anschluss wird zusätzlich ein Hintergrundgeräusch abgespielt. Stellen Sie nun die Lautstärke des Hintergrundgeräusches so laut ein, wie Sie es gerade noch ertragen würden. Sie sollten den Sätzen in dieser Situation über eine längere Zeit folgen können. Das Zuhören sollte Sie nicht ermüden oder anstrengen. Zuerst stellen Sie dazu bitte das Hintergrundgeräusch zu laut ein und danach leiser bis die Sätze sehr klar und deutlich zu verstehen sind. Zum Schluss regeln Sie den maximalen Pegel ein, den Sie über einen langen Zeitraum akzeptieren würden, während Sie den Sätzen zuhören.



Anleitung:

Die Wiedergabe der Sätze wird mit der Start-Taste begonnen. Die Sätze werden über einen Kopfhörer dargeboten. Nachdem Sie den Sätzen eine Weile zugehört haben, sollten Sie die Lautstärke mit Hilfe der oben abgebildeten Oberfläche einstellen. Sie können die Lautstärke regeln, indem Sie die Pfeiltasten (Pfeil oben, + => lauter; Pfeil unten, - => leiser) verwenden. Die Meldungen im Textfeld weisen Sie darauf hin, wie Sie die

Lautstärke einzustellen haben. Jede Einstellung muss mit der OK-Taste bestätigt werden.

Verändern Sie bitte die **Lautstärke, bis die Sätze** im Vergleich zu einer angenehmen Lautstärke **zu laut** sind. Danach verringern Sie bitte die Lautstärke wieder, bis die Sätze im Vergleich zu einer angenehmen Lautstärke **zu leise** sind. Abschließend erhöhen Sie bitte die Lautstärke der Sätze, bis sie **sehr angenehm** ist. Stellen Sie also Ihre eigene bevorzugte Lautstärke ein. Jede Einstellung der Lautstärke bestätigen Sie bitte mit der OK-Taste.

Anschließend wird ein **Hintergrundgeräusch** zusätzlich dargeboten. Verändern Sie seine Lautstärke, bis die **Sätze nicht mehr zu verstehen** sind. Anschließend verringern Sie die Lautstärke des Hintergrundgeräusches, bis die **Sätze sehr klar und deutlich zu verstehen** sind. Sie sollten den Sätzen leicht folgen können. Erhöhen Sie nun die Lautstärke des Hintergrundgeräusches auf eine für Sie **maximal tolerierbare Lautstärke**. Das Hintergrundgeräusch sollte so laut sein, dass Sie den Sätzen über einen längeren Zeitraum zuhören können ohne sich anzustrengen oder zu ermüden. Bitte bestätigen Sie auch hier jede Einstellung der Lautstärke mit der OK-Taste.

Vielen Dank!

Instruktionen, die im Test auf dem Bildschirm dargestellt werden:

- Verändern Sie die Lautstärke bis die **Sätze** im Vergleich zu einer angenehmen Lautstärke **zu laut** sind. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Verringern Sie die Lautstärke wieder, bis die **Sätze** im Vergleich zu einer angenehmen Lautstärke **zu leise** sind. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Erhöhen Sie die Lautstärke der **Sätze**, bis sie **sehr angenehm** ist. Stellen Sie also ihre eigene **bevorzugte Lautstärke** ein. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Verändern Sie die Lautstärke des **Hintergrundgeräusches**, bis die **Sätze nicht mehr zu verstehen** sind. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Verringern Sie die Lautstärke des **Hintergrundgeräusches**, bis die **Sätze sehr klar und deutlich zu verstehen** sind. Sie sollten den Sätzen leicht folgen können. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Erhöhen Sie nun die Lautstärke des **Hintergrundgeräusches** auf eine für Sie **maximal tolerierbare Lautstärke**. Das Hintergrundgeräusch sollte so laut sein, dass Sie den Sätzen über einen längeren Zeitraum zuhören können ohne sich anzustrengen oder zu ermüden. Beginnen Sie die Wiedergabe der Sätze mit „Start“.

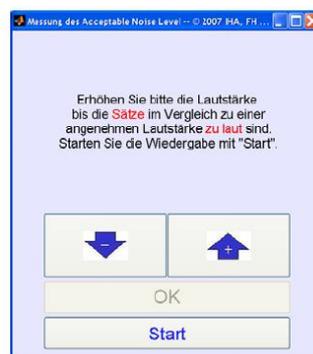
B.2 Verwendete Instruktionen bei der Untersuchung ihres Einflusses

B.2.1 Instruktion: Aufsteigend

Annehmbarer Pegel des Hintergrundgeräusches

Ihnen werden gleich Sätze dargeboten, die von einem Sprecher vorgelesen werden. Als Erstes werden Sie gebeten, die Lautstärke der Sätze auf einen angenehmen Pegel einzustellen, d. h. etwa so, als ob Sie Radio hören würden. Dazu soll die Lautstärke erst lauter und danach leiser eingestellt werden. Danach wählen Sie bitte eine Lautstärke die Ihnen sehr angenehm ist.

Im Anschluss wird zusätzlich ein Hintergrundgeräusch abgespielt. Stellen Sie nun die Lautstärke des Hintergrundgeräusches so laut ein, wie Sie es gerade noch ertragen würden. Sie sollten den Sätzen in dieser Situation über eine längere Zeit folgen können. Das Zuhören sollte Sie nicht ermüden oder anstrengen. Zuerst stellen Sie dazu bitte das Hintergrundgeräusch zu laut ein und danach leiser bis die Sätze sehr klar und deutlich zu verstehen sind. Zum Schluss regeln Sie den maximalen Pegel ein, den Sie über einen langen Zeitraum akzeptieren würden, während Sie den Sätzen zuhören.



Anleitung:

Die Wiedergabe der Sätze wird mit der Start-Taste begonnen. Die Sätze werden über einen Kopfhörer dargeboten. Nachdem Sie den Sätzen eine Weile zugehört haben, sollten Sie die Lautstärke mit Hilfe der oben abgebildeten Oberfläche einstellen. Sie können die Lautstärke regeln, indem Sie die Pfeiltasten (Pfeil oben, + => lauter; Pfeil unten, - => leiser) verwenden. Die Meldungen im Textfeld weisen Sie darauf hin, wie Sie die Lautstärke einzustellen haben. Jede Einstellung muss mit der OK-Taste bestätigt werden.

Verändern Sie bitte die **Lautstärke, bis die Sätze** im Vergleich zu einer angenehmen Lautstärke **zu laut** sind. Danach verringern Sie bitte die Lautstärke wieder, bis die Sätze im Vergleich zu einer angenehmen Lautstärke **zu leise** sind. Abschließend erhöhen Sie bitte die Lautstärke der Sätze, bis sie **sehr angenehm** ist. Stellen Sie also Ihre eigene bevorzugte Lautstärke ein. Jede Einstellung der Lautstärke bestätigen Sie bitte mit der OK-Taste.

Anschließend wird ein **Hintergrundgeräusch** zusätzlich dargeboten. Verändern Sie seine Lautstärke, bis die **Sätze nicht mehr zu verstehen** sind. Anschließend ver-

ringern Sie die Lautstärke des Hintergrundgeräusches, bis die **Sätze sehr klar und deutlich zu verstehen** sind. Sie sollten den Sätzen leicht folgen können. Erhöhen Sie nun die Lautstärke des Hintergrundgeräusches auf eine für Sie **maximal tolerierbare Lautstärke**. Das Hintergrundgeräusch sollte so laut sein, dass Sie den Sätzen über einen längeren Zeitraum zuhören können ohne sich anzustrengen oder zu ermüden. Bitte bestätigen Sie auch hier jede Einstellung der Lautstärke mit der OK-Taste.

Vielen Dank!

Instruktionen, die im Test auf dem Bildschirm dargestellt werden:

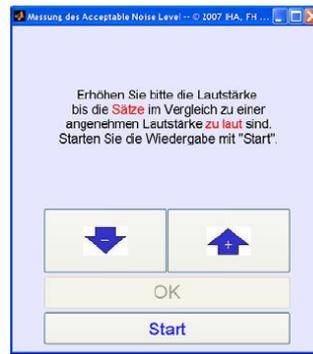
- Verändern Sie die Lautstärke bis die **Sätze** im Vergleich zu einer angenehmen Lautstärke **zu laut** sind. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Verringern Sie die Lautstärke wieder, bis die **Sätze** im Vergleich zu einer angenehmen Lautstärke **zu leise** sind. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Erhöhen Sie die Lautstärke der **Sätze**, bis sie **sehr angenehm** ist. Stellen Sie also ihre eigene **bevorzugte Lautstärke** ein. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Verändern Sie die Lautstärke des **Hintergrundgeräusches**, bis die **Sätze nicht mehr zu verstehen** sind. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Verringern Sie die Lautstärke des **Hintergrundgeräusches**, bis die **Sätze sehr klar und deutlich zu verstehen** sind. Sie sollten den Sätzen leicht folgen können. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Erhöhen Sie nun die Lautstärke des **Hintergrundgeräusches** auf eine für Sie **maximal tolerierbare Lautstärke**. Das Hintergrundgeräusch sollte so laut sein, dass Sie den Sätzen über einen längeren Zeitraum zuhören können ohne sich anzustrengen oder zu ermüden. Beginnen Sie die Wiedergabe der Sätze mit „Start“.

B.2.2 Instruktion: Absteigend

Annehmbarer Pegel des Hintergrundgeräusches

Ihnen werden gleich Sätze dargeboten, die von einem Sprecher vorgelesen werden. Als Erstes werden Sie gebeten, die Lautstärke der Sätze auf einen angenehmen Pegel einzustellen, d. h. etwa so, als ob Sie Radio hören würden. Dazu soll die Lautstärke erst leiser und danach lauter eingestellt werden. Danach wählen Sie bitte eine Lautstärke die Ihnen sehr angenehm ist.

Im Anschluss wird zusätzlich ein Hintergrundgeräusch abgespielt. Stellen Sie nun die Lautstärke des Hintergrundgeräusches so laut ein, wie Sie es gerade noch ertragen würden. Sie sollten den Sätzen in dieser Situation über eine längere Zeit folgen können. Das Zuhören sollte Sie nicht ermüden oder anstrengen. Zuerst stellen Sie dazu bitte das Hintergrundgeräusch zu leise ein und danach lauter bis die Sätze nicht mehr zu verstehen sind. Zum Schluss regeln Sie den maximalen Pegel ein, den Sie über einen langen Zeitraum akzeptieren würden, während Sie den Sätzen zuhören.



Anleitung:

Die Wiedergabe der Sätze wird mit der Start-Taste begonnen. Die Sätze werden über einen Kopfhörer dargeboten. Nachdem Sie den Sätzen eine Weile zugehört haben, sollten Sie die Lautstärke mit Hilfe der oben abgebildeten Oberfläche einstellen. Sie können die Lautstärke regeln, indem Sie die Pfeiltasten (Pfeil oben, + => lauter; Pfeil unten, - => leiser) verwenden. Die Meldungen im Textfeld weisen Sie darauf hin, wie Sie die Lautstärke einzustellen haben. Jede Einstellung muss mit der OK-Taste bestätigt werden.

Verändern Sie bitte die **Lautstärke, bis die Sätze** im Vergleich zu einer angenehmen Lautstärke **zu leise** sind. Danach erhöhen Sie bitte die Lautstärke wieder, bis die Sätze im Vergleich zu einer angenehmen Lautstärke **zu laut** sind. Abschließend verringern Sie bitte die Lautstärke der Sätze, bis sie **sehr angenehm** ist. Stellen Sie also Ihre eigene bevorzugte Lautstärke ein. Jede Einstellung der Lautstärke bestätigen Sie bitte mit der OK-Taste.

Anschließend wird ein **Hintergrundgeräusch** zusätzlich dargeboten. Verändern Sie seine Lautstärke, bis die **Sätze sehr klar und deutlich zu verstehen** sind. Sie sollten den Sätzen leicht folgen können. Anschließend erhöhen Sie die Lautstärke des Hintergrundgeräusches, bis die **Sätze nicht mehr zu verstehen** sind. Verringern Sie nun die Lautstärke des Hintergrundgeräusches auf eine für Sie **maximal tolerierbare Lautstärke**. Das Hintergrundgeräusch sollte so laut sein, dass Sie den Sätzen über einen längeren Zeitraum zuhören können ohne sich anzustrengen oder zu ermüden. Bitte bestätigen Sie auch hier jede Einstellung der Lautstärke mit der OK-Taste.

Vielen Dank!

Instruktionen, die im Test auf dem Bildschirm dargestellt werden:

- Verändern Sie die Lautstärke, bis die **Sätze** im Vergleich zu einer angenehmen Lautstärke **zu leise** sind. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Erhöhen Sie die Lautstärke bis die **Sätze** im Vergleich zu einer angenehmen Lautstärke **zu laut** sind. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Verringern Sie die Lautstärke der **Sätze**, bis sie **sehr angenehm** ist. Stellen Sie also ihre **eigene bevorzugte Lautstärke** ein. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Verändern Sie die Lautstärke des **Hintergrundgeräusches**, bis die **Sätze sehr klar und deutlich zu verstehen** sind. Sie sollten den Sätzen leicht folgen können.

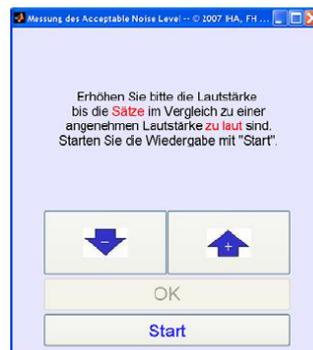
nen. Beginnen Sie die Wiedergabe der Sätze mit „Start“.

- Erhöhen Sie die Lautstärke des **Hintergrundgeräusches**, bis die **Sätze nicht mehr zu verstehen** sind. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Verringern Sie nun die Lautstärke des **Hintergrundgeräusches** auf eine für Sie **maximal tolerierbare Lautstärke**. Das Hintergrundgeräusch sollte so laut sein, dass Sie den Sätzen über einen längeren Zeitraum zuhören können ohne sich anzustrengen oder zu ermüden. Beginnen Sie die Wiedergabe der Sätze mit „Start“.

B.3 Instruktion der MCL-Messung

Annehmbarer Pegel der Sprache

Ihnen werden gleich Sätze dargeboten, die von einem Sprecher vorgelesen werden. Als Erstes werden Sie gebeten, die Lautstärke der Sätze auf einen angenehmen Pegel einzustellen, d. h. etwa so, als ob Sie Radio hören würden. Dazu soll die Lautstärke erst lauter und danach leiser eingestellt werden. Danach wählen Sie bitte eine Lautstärke die Ihnen sehr angenehm ist.



Anleitung:

Die Wiedergabe der Sätze wird mit der Start-Taste begonnen. Die Sätze werden über einen Kopfhörer dargeboten. Nachdem Sie den Sätzen eine Weile zugehört haben, sollten Sie die Lautstärke mit Hilfe der oben abgebildeten Oberfläche einstellen. Sie können die Lautstärke regeln, indem Sie die Pfeiltasten (Pfeil oben, + => lauter; Pfeil unten, - => leiser) verwenden. Die Meldungen im Textfeld weisen Sie darauf hin, wie Sie die Lautstärke einzustellen haben. Jede Einstellung muss mit der OK-Taste bestätigt werden.

Erhöhen Sie bitte die **Lautstärke**, bis die **Sätze** im Vergleich zu einer angenehmen Lautstärke **zu laut** sind. Danach verringern Sie bitte die Lautstärke wieder, bis die Sätze im Vergleich zu einer angenehmen Lautstärke **zu leise** sind. Abschließend erhöhen Sie bitte die Lautstärke der Sätze, bis sie **sehr angenehm** ist. Stellen Sie also Ihre eigene bevorzugte Lautstärke ein. Jede Einstellung der Lautstärke bestätigen Sie bitte mit der OK-Taste.

Vielen Dank!

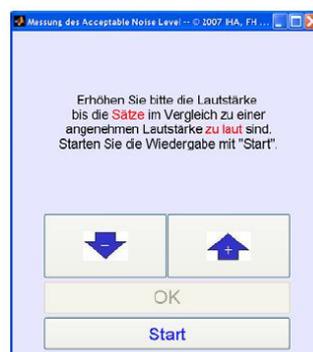
Instruktionen, die im Test auf dem Bildschirm dargestellt werden:

- Verändern Sie die Lautstärke, bis die **Sätze** im Vergleich zu einer angenehmen Lautstärke **zu laut** sind. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Verringern Sie die Lautstärke wieder, bis die **Sätze** im Vergleich zu einer angenehmen Lautstärke **zu leise** sind. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Erhöhen Sie die Lautstärke der **Sätze**, bis sie **sehr angenehm** ist. Stellen Sie also ihre eigene **bevorzugte Lautstärke** ein. Beginnen Sie die Wiedergabe der Sätze mit „Start“.

B.4 Instruktion der BNL-Messung

Annehmbarer Pegel des Hintergrundgeräusches

Ihnen werden gleich Sätze dargeboten, die von einem Sprecher vorgelesen werden. Sie werden zusammen mit einem Hintergrundgeräusch abgespielt. Stellen Sie die Lautstärke des Hintergrundgeräusches so laut ein, wie Sie es gerade noch ertragen würden. Sie sollten den Sätzen in dieser Situation über eine längere Zeit folgen können. Das Zuhören sollte Sie nicht ermüden oder anstrengen. Zuerst stellen Sie dazu bitte das Hintergrundgeräusch zu laut ein und danach leiser bis die Sätze sehr klar und deutlich zu verstehen sind. Zum Schluss regeln Sie den maximalen Pegel ein, den Sie über einen langen Zeitraum akzeptieren würden, während Sie den Sätzen zuhören.



Anleitung:

Die Wiedergabe der Sätze und des Hintergrundgeräusches wird mit der Start-Taste begonnen. Die Signale werden über einen Kopfhörer dargeboten. Nachdem Sie den Sätzen eine Weile zugehört haben, sollten Sie die Lautstärke des Hintergrundgeräusches mit Hilfe der oben abgebildeten Oberfläche einstellen. Sie können die Lautstärke regeln, indem Sie die Pfeiltasten (Pfeil oben, + => lauter; Pfeil unten, - => leiser) verwenden. Die Meldungen im Textfeld weisen Sie darauf hin, wie Sie die Lautstärke einzustellen haben. Jede Einstellung muss mit der OK-Taste bestätigt werden.

Erhöhen Sie die Lautstärke des Hintergrundgeräusches, bis die **Sätze nicht mehr zu verstehen** sind. Anschließend verringern Sie die Lautstärke des Hintergrundgeräusches, bis die **Sätze sehr klar und deutlich zu verstehen** sind. Sie sollten den Sätzen

leicht folgen können. Erhöhen Sie nun die Lautstärke des Hintergrundgeräusches auf eine für Sie **maximal tolerierbare Lautstärke**. Das Hintergrundgeräusch sollte so laut sein, dass Sie den Sätzen über einen längeren Zeitraum zuhören können ohne sich anzustrengen oder zu ermüden. Bitte bestätigen Sie jede Einstellung der Lautstärke mit der OK-Taste.

Vielen Dank!

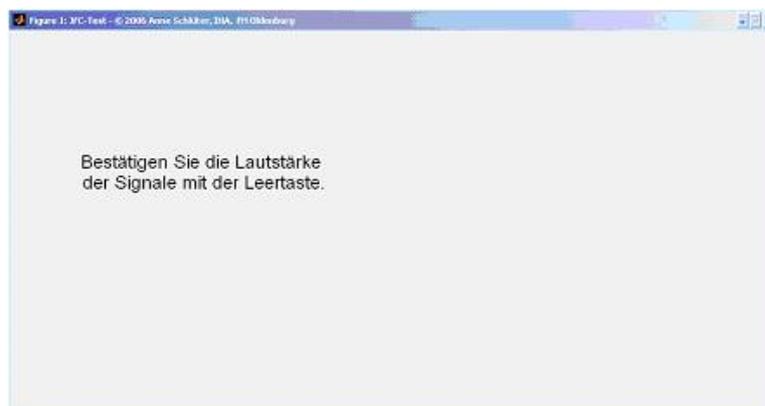
Instruktionen, die im Test auf dem Bildschirm dargestellt werden:

- Verändern Sie die Lautstärke des **Hintergrundgeräusches**, bis die **Sätze nicht mehr zu verstehen** sind. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Verringern Sie die Lautstärke des **Hintergrundgeräusches**, bis die **Sätze sehr klar und deutlich zu verstehen** sind. Sie sollten den Sätzen leicht folgen können. Beginnen Sie die Wiedergabe der Sätze mit „Start“.
- Erhöhen Sie nun die Lautstärke des **Hintergrundgeräusches** auf eine für Sie **maximal tolerierbare Lautstärke**. Das Hintergrundgeräusch sollte so laut sein, dass Sie den Sätzen über einen längeren Zeitraum zuhören können ohne sich anzustrengen oder zu ermüden. Beginnen Sie die Wiedergabe der Sätze mit „Start“.

B.5 Instruktion des Just Follow Conversation Tests

Subjektive Sprachverständlichkeit

Ihnen werden gleich Sätze zusammen mit einem Hintergrundgeräusch dargeboten. Ihre Aufgabe ist, die Lautstärke des Rauschens zu regeln. Dabei sollten sie einen Pegel erreichen, bei dem Sie die Sätze gerade eben verstehen können.



Anleitung:

Die Sätze und das Hintergrundgeräusch werden über einen Kopfhörer dargeboten. Die Signalwiedergabe starten Sie durch Betätigung der linken Maustaste. Sie sollten die Lautstärke des Hintergrundgeräusches mit Hilfe der Maus einstellen. Bewegen Sie den roten Ball von sich weg, wird das Hintergrundgeräusch lauter. Bewegen Sie den roten Ball zu sich hin, wird das Hintergrundgeräusch leiser. Bitte stellen Sie so eine Lautstärke des Hintergrundgeräusches ein, bei der Sie die Sätze gerade eben verstehen können.

Haben Sie die gewünschte Lautstärke eingestellt, warten Sie bitte, bis im Textfeld auf dem Bildschirm eine Aufforderung erscheint. Dann können Sie die Lautstärke mit der Leertaste auf der Tastatur bestätigen.

Dieser Vorgang wiederholt sich zwölfmal. Das Rauschen wird zu Beginn jeder Messung lauter oder leiser eingestellt. Bitte stellen Sie bei jeder Messung die Lautstärke des Hintergrundgeräusches ein, bei der Sie die Sätze gerade eben verstehen.

Vielen Dank!

B.6 Instruktion des Oldenburger Satztests

Sprachverständlichkeitsmessung

Dies ist ein Test, um festzustellen, wie gut Sie Sprache in geräuschvoller Umgebung verstehen können. Hierzu werden Ihnen Sätze dargeboten, die von einer männlichen Stimme gesprochen werden. Jeder Satz besteht aus 5 Wörtern der Struktur: Name Verb Zahl Adjektiv Objekt. (z.B. Ulrich schenkt sieben schwere Sessel.)

Die Sätze sind nicht unbedingt sinnvoll. Zusätzlich zu der Sprache wird ein Rauschen dargeboten.

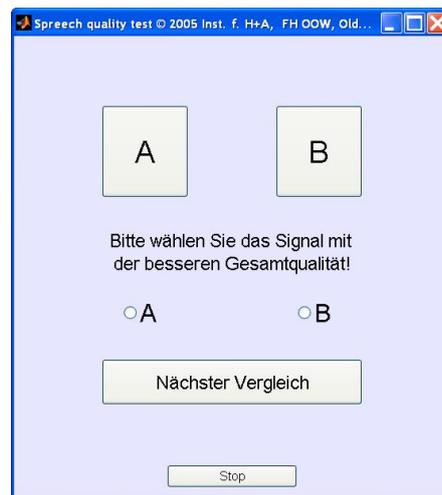
Bitte wiederholen Sie nach der Darbietung den Satz oder jedes Wort, welches Sie verstanden haben. Wenn Sie unsicher sind, dürfen Sie gerne auch raten. Während der Messung wird das Rauschen in der Lautstärke verändert. Die Sprache kann dadurch teilweise sehr leise sein. Lassen Sie sich dadurch bitte nicht entmutigen. Für die Aussagekraft der Messung ist es wichtig, diese Messung unter schwierigen Bedingungen durchzuführen.

Vielen Dank!

B.7 Instruktion des Paarvergleichs

Paarvergleich

Ihnen werden gleich zwei identische Sätze mit einem Hintergrundgeräusch dargeboten. Diese Klangbeispiele sind unterschiedlich verarbeitet worden. Bitte wählen Sie das Klangbeispiel, das eine bessere Gesamtqualität besitzt.



Anleitung:

Die zwei Klangbeispiele werden über einen Kopfhörer dargeboten. Sie sind benannt mit den Buchstaben „A“ und „B“. Die Signalwiedergabe starten Sie durch Betätigung des Feldes „A“ und/oder „B“. Bitte hören Sie sich beide Klangbeispiele an. Danach wählen Sie auf der dargestellten Oberfläche das Signal mit der besseren Gesamtqualität und bestätigen Ihre Wahl mit „Nächster Vergleich“. Anschließend sind die Felder „A“ und „B“ mit neuen Klangbeispielen hinterlegt, deren Gesamtqualität Sie bitte ebenfalls bewerten.

Vielen Dank!

Anhang C

Hörverluste der Versuchspersonen

C.1 Untersuchung der Schrittgröße

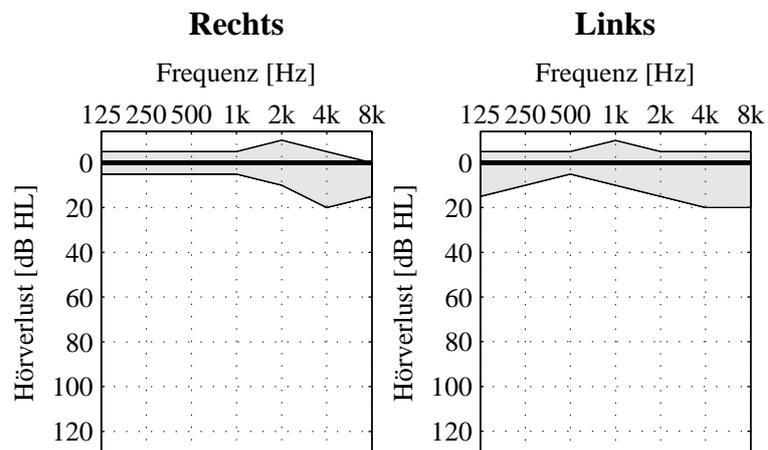


Abb. C.1: Spanne des Hörvermögens der Versuchspersonen, die bei der Untersuchung der Schrittgröße im ANLT teilgenommen haben, dargestellt in einem Tonaudiogramm.

C.2 Untersuchung einer auf- und absteigenden Instruktion

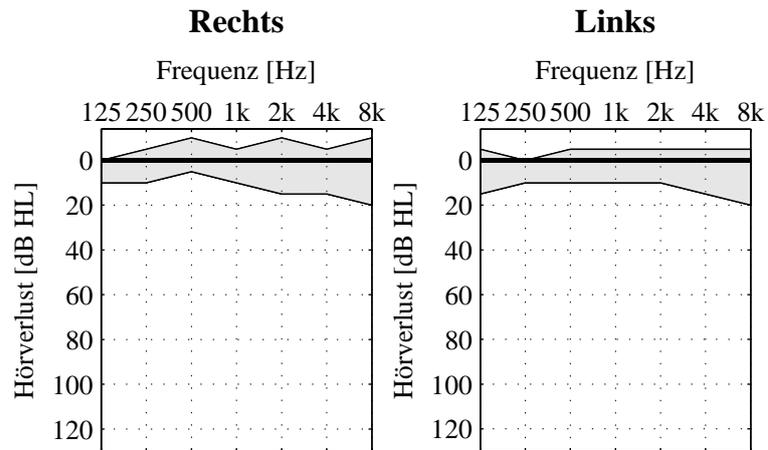


Abb. C.2: Spanne des Hörvermögens der Versuchspersonen, die bei der Untersuchung der Instruktion des ANLT teilgenommen haben, dargestellt in einem Tonaudiogramm.

C.3 Untersuchung des Startpegels

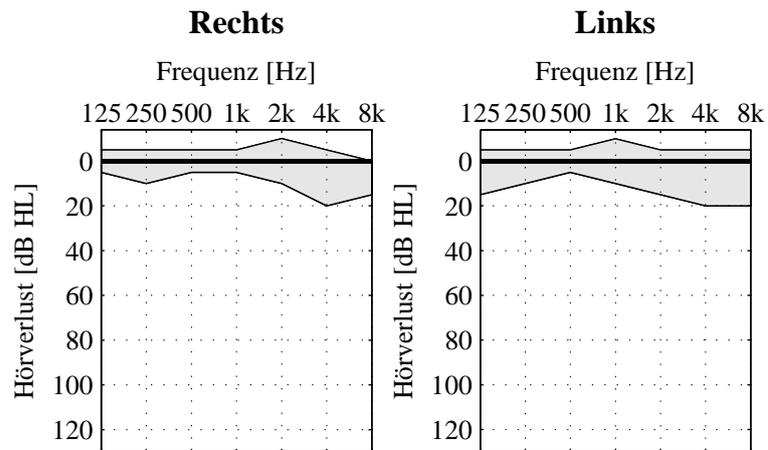


Abb. C.3: Spanne des Hörvermögens der Versuchspersonen, die bei der Untersuchung des Startpegels im ANLT teilgenommen haben, dargestellt in einem Tonaudiogramm.

C.4 Untersuchung der Test-Retest-Reliabilität und Reliabilität über einen kurzen Zeitraum

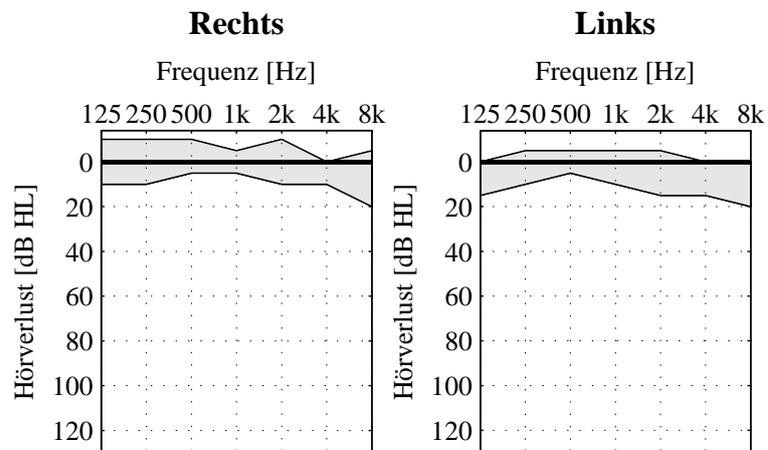


Abb. C.4: Spanne des Hörvermögens der Versuchspersonen, die bei der Untersuchung Test-Retest-Reliabilität und der Reliabilität über einen kurzen Zeitraum teilgenommen haben, dargestellt in einem Tonaudiogramm.

C.5 Einfluss der Methodik bei Wiederholungsmessungen auf den ANL

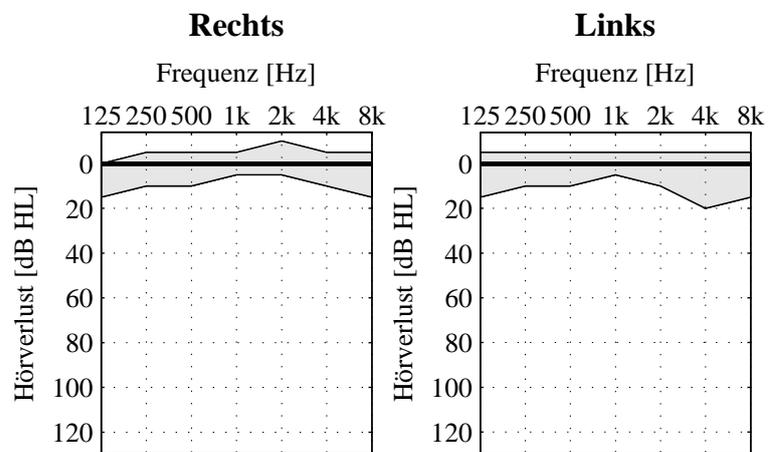


Abb. C.5: Spanne des Hörvermögens der Versuchspersonen, die bei der Untersuchung der Methodik bei Wiederholungsmessungen auf den ANL teilgenommen haben, dargestellt in einem Tonaudiogramm.

C.6 Einfluss einer SRT-Messung auf ANL- und JFC-Werte

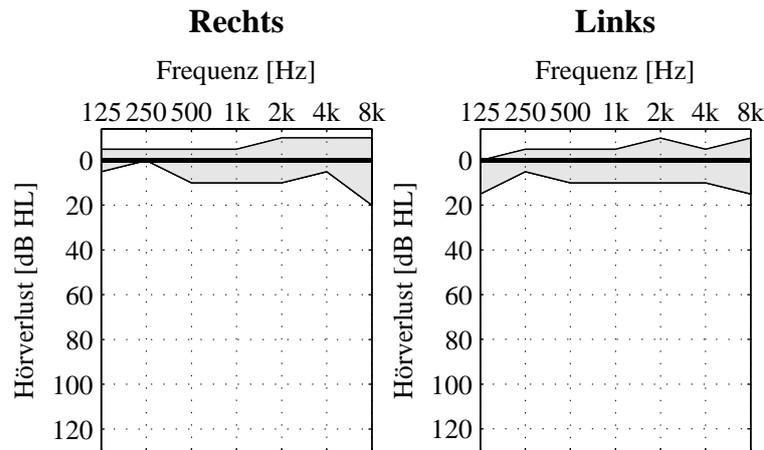


Abb. C.6: Spanne des Hörvermögens der Versuchspersonen, die bei der Untersuchung des Einflusses einer SRT-Messung auf ANL- und JFC-Werte teilgenommen haben, dargestellt in einem Tonaudiogramm.

C.7 Einfluss unterschiedlicher Hintergrundgeräusche

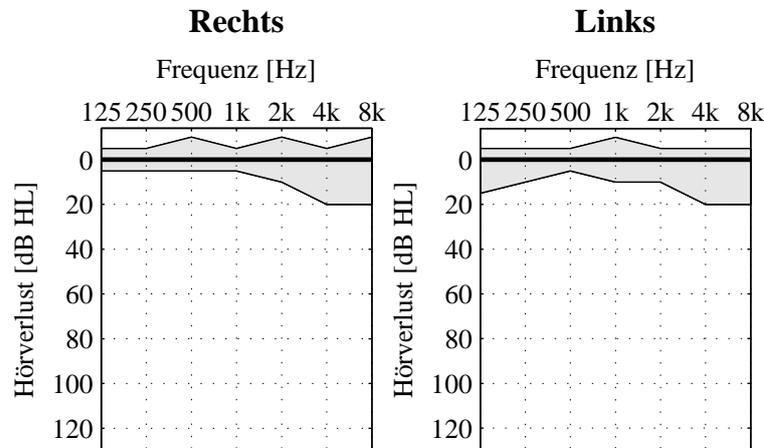


Abb. C.7: Spanne des Hörvermögens der Versuchspersonen, die bei der Untersuchung unterschiedlicher Hintergrundgeräusches teilgenommen haben, dargestellt in einem Tonaudiogramm.

C.8 Einfluss des Darbietungspegels bei der Übung des Sprachmaterials

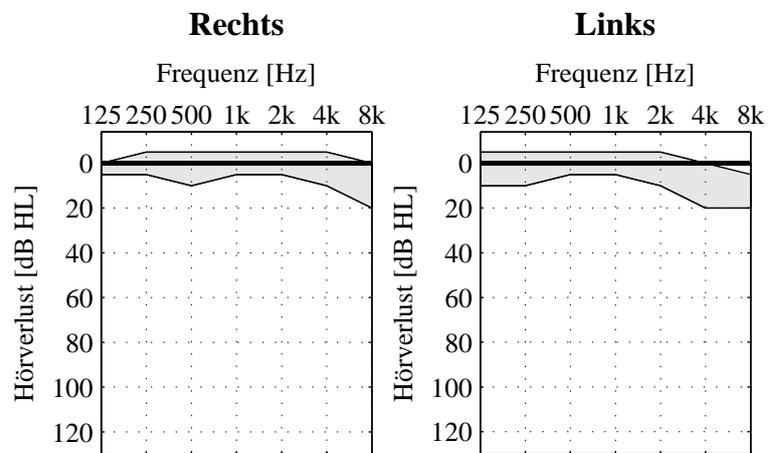


Abb. C.8: Spanne des Hörvermögens der Versuchspersonen, die bei der Untersuchung des Einflusses des Darbietungspegels bei der Übung des Sprachmaterials teilgenommen haben, dargestellt in einem Tonaudiogramm.

Anhang D

Untersuchung des ANLT

D.1 Experiment A: Untersuchung der Schrittgröße

D.1.1 Signifikanzprüfung der Messwerte im ANLT

Tab. D.1: χ^2 -Werte und Irrtumswahrscheinlichkeit p eines Friedman-Tests der Messwerte im ANLT. Die Messwerte wurden mit unterschiedlichen Schrittgrößen ermittelt.

| Messwert im ANLT | $\chi^2_{(2;20;0,05)}$ | P |
|--------------------|------------------------|-------|
| maximale Sprache | 11,8429 | 0,003 |
| minimale Sprache | 3,25 | 0,197 |
| MCL | 4,439 | 0,109 |
| maximales Rauschen | 11,436 | 0,003 |
| minimales Rauschen | 3,706 | 0,157 |
| BNL | 2,513 | 0,285 |
| ANL | 2,722 | 0,256 |

D.1.2 ANL-Werte gemessen mit unterschiedlichen Schrittgrößen

In Abb. D.1 sind die ANL-Werte jeder Versuchsperson bei jeder Schrittgröße aufgetragen. Die Daten haben einen großen Streubereich, d. h., die ermittelten Messwerte unterscheiden sich deutlich zwischen den Versuchspersonen.

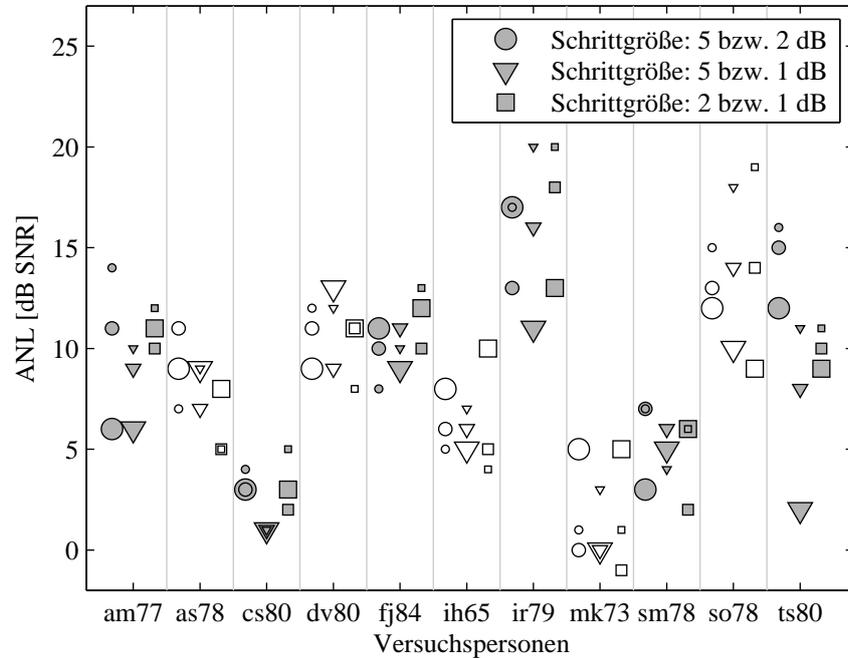


Abb. D.1: Darstellung der eingestellten ANL-Werte jeder Versuchsperson bei unterschiedlichen Schrittgrößen. Große Symbole stehen für die erste, mittlere Symbole für die zweite und kleine Symbole für die dritte Messung.

D.2 Experiment B: Untersuchung einer auf- oder absteigenden Instruktion

D.2.1 Messwerte des ANLT, gemessen mit auf- und absteigender Instruktion

Abb. D.2 zeigt einen Boxplot der einzelnen Messwerte des ANLT, die mit den Instruktionen „Aufsteigend“ und „Absteigend“ erreicht wurden. Die Ergebnisse sind in der Reihenfolge dargestellt, in der sie gemessen bzw. instruiert wurden. Deutlich zu erkennen ist, dass mit der Instruktion „Aufsteigend“ zuerst ein hoher, danach ein niedriger und anschließend ein mittlerer Pegel der Sprache und des Rauschen eingestellt wurde. Mit der Instruktion „Absteigend“ verhalten sich die eingestellten Pegel konträr. Sie liegen bei niedrigen, hohen und mittleren Werten.

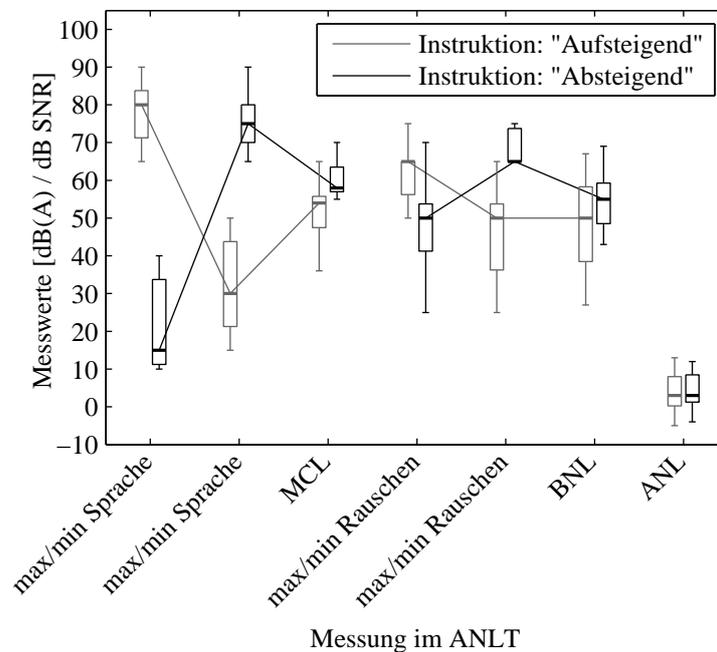


Abb. D.2: Messwerte im ANLT, gemessen mit auf- oder absteigender Instruktion. Dargestellt sind die Ergebnisse in der Reihenfolge in der sie gemessen wurden. Die Ergebnisse maximale Sprache, minimale Sprache, MCL, maximales Rauschen, minimales Rauschen und BNL sind in dB(A), die Werte des ANL in dB SNR angegeben.

D.2.2 Signifikanzprüfung der Messwerte im ANLT

Tab. D.2: Irrtumswahrscheinlichkeit p eines Wilcoxon Vorzeichen-Rang-Tests der Messwerte im ANLT. Die Messwerte wurden mit auf- und absteigender Instruktion ermittelt.

| Messwert im ANLT | p |
|--------------------|-------|
| maximale Sprache | 0,109 |
| minimale Sprache | 0,018 |
| MCL | 0,003 |
| maximales Rauschen | 0,078 |
| minimales Rauschen | 0,344 |
| BNL | 0,007 |
| ANL | 0,898 |

D.3 Experiment C: Untersuchung des Startpegels

Tab. D.3: χ^2 -Werte und Irrtumswahrscheinlichkeit p eines Friedman-Tests der Messwerte im ANLT. Die Messwerte wurden mit unterschiedlichen Startpegeln ermittelt.

| Messwert im ANLT | $\chi^2_{(2;22;0,05)}$ | p |
|--------------------|------------------------|-------|
| maximale Sprache | 6,737 | 0,034 |
| minimale Sprache | 7,807 | 0,02 |
| MCL | 10,978 | 0,004 |
| maximales Rauschen | 6,75 | 0,034 |
| minimales Rauschen | 4,57 | 0,102 |
| BNL | 3,591 | 0,102 |
| ANL | 7,143 | 0,028 |

D.4 Experiment D: Test-Retest-Reliabilität und Reliabilität über einen kurzen Zeitraum

D.4.1 Gegenüberstellung der Messwerte des ANLT des Termin A und B

Zur Untersuchung der Test-Retest-Reliabilität wurde für jede Versuchsperson der Median der zwölf Messwiederholungen der Termine A und B ermittelt. Diese Werte des MCL, BNL und ANL sind in den Teilabbildungen D.3a, D.3b und D.3c dargestellt und liegen nahe der Winkelhalbierenden.

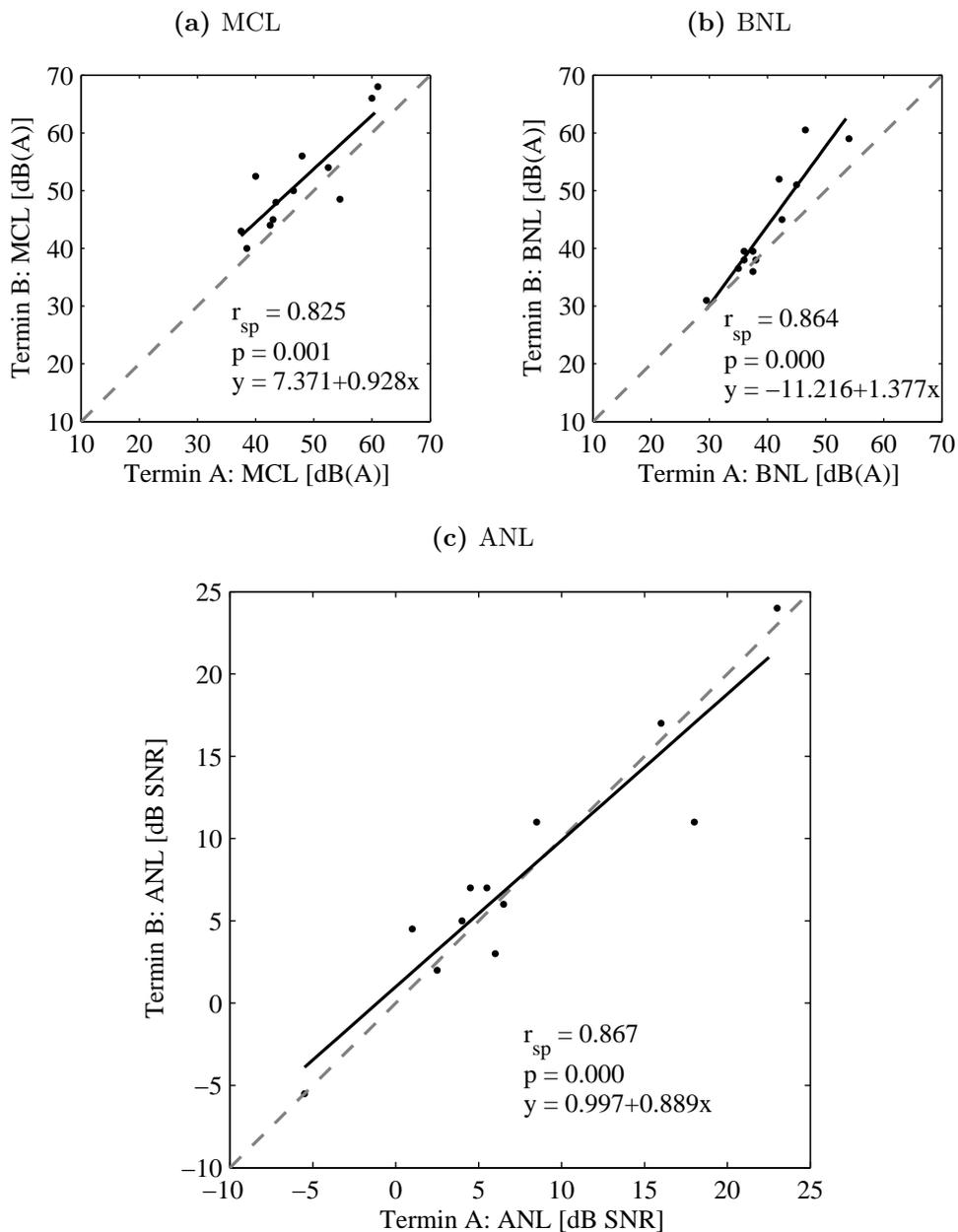


Abb. D.3: Gegenüberstellung der Werte des (a) MCL , (b) BNL und (c) ANL, gemessen in den Terminen A und B.

D.4.2 Inter- und intraindividuelle Streuung des ANL

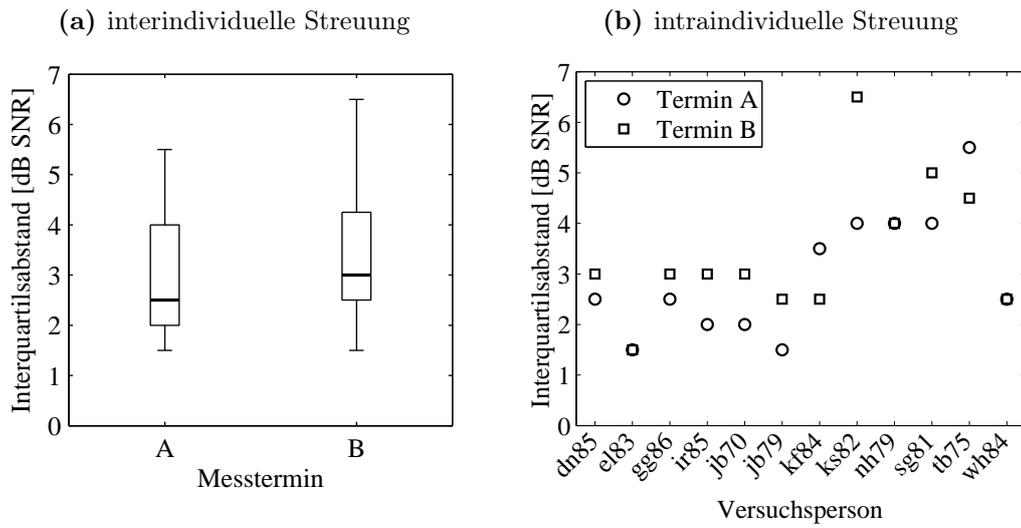


Abb. D.4: (a) interindividuelle Streuung und (b) die intraindividuelle Streuung der ANL-Werte, gemessen in den Terminen A und B.

D.4.3 Signifikanzprüfung der kumulierten ANL-Werte

Tab. D.4: Prüfung der signifikanten Unterschiede zwischen den kumulierten ANL-Werten des Kap. 3.7 mit einem Friedman-Test.

| Verglichene kumulierte Anzahl | $\chi^2_{(1;24;0,05)}$ | Irrtumswahrscheinlichkeit p |
|-------------------------------|------------------------|-------------------------------|
| 1 und 3 | 4,469 | 0,035 |
| 1 und 5 | 4,673 | 0,031 |
| 1 und 6 | 7,022 | 0,008 |
| 1 und 7 | 6,182 | 0,013 |
| 1 und 8 | 9,677 | 0,002 |
| 1 und 9 | 10,889 | 0,001 |
| 1 und 10 | 12,162 | 0,001 |
| 1 und 11 | 12,162 | 0,001 |
| 1 und 12 | 14,204 | 0,000 |
| 2 und 9 | 3,977 | 0,046 |
| 2 und 10 | 6,368 | 0,012 |
| 2 und 11 | 7,149 | 0,008 |
| 2 und 12 | 9,218 | 0,002 |
| 3 und 9 | 4,335 | 0,037 |
| 3 und 10 | 5,207 | 0,023 |
| 3 und 11 | 4,765 | 0,029 |
| 3 und 12 | 5,505 | 0,019 |
| 4 und 9 | 6,541 | 0,011 |
| 4 und 10 | 7,714 | 0,006 |
| 4 und 11 | 8,297 | 0,004 |
| 4 und 12 | 9,677 | 0,002 |
| 5 und 9 | 5,714 | 0,017 |
| 5 und 10 | 6,661 | 0,010 |
| 5 und 11 | 6,849 | 0,009 |
| 5 und 12 | 8,895 | 0,003 |
| 6 und 9 | 4,500 | 0,034 |
| 6 und 10 | 6,014 | 0,014 |
| 6 und 11 | 7,214 | 0,007 |
| 6 und 12 | 8,152 | 0,004 |
| 7 und 10 | 4,629 | 0,031 |
| 7 und 11 | 4,292 | 0,038 |
| 7 und 12 | 7,085 | 0,008 |
| 8 und 10 | 4,250 | 0,039 |
| 8 und 11 | 3,840 | 0,050 |
| 8 und 12 | 6,014 | 0,014 |

D.4.4 Berücksichtigung der interindividuellen Streuung im kumulierten ANL

Um die Ergebnisse des ANLT der Abb. 3.6c und 3.7 in Kap. 3.7 unabhängig von der interindividuellen Streuung zu betrachten, wurde von den Ergebnissen der zwölf Wiederholungsmessungen jeder Versuchsperson der erste ermittelte ANL subtrahiert. Die Ergebnisse sind in Abb. D.5 dargestellt. Abb. D.5a zeigt die ANL-Werte der Abb. 3.6c, nachdem der erste ANL-Wert subtrahiert wurde. Für Abb. D.5b wurden dann die Mediane aus der auf der Abszisse angegebenen Anzahl der ANL-Werte der Abb. D.5a ermittelt. Die Bildung der Mediane wurde wie für Abb. 3.7 vorgenommen. In diesen Abbildungen ist deutlich die geringe Zunahme des medianen ANL um ca. 2 dB SNR durch eine Anzahl von zwölf Messwiederholungen zu erkennen.

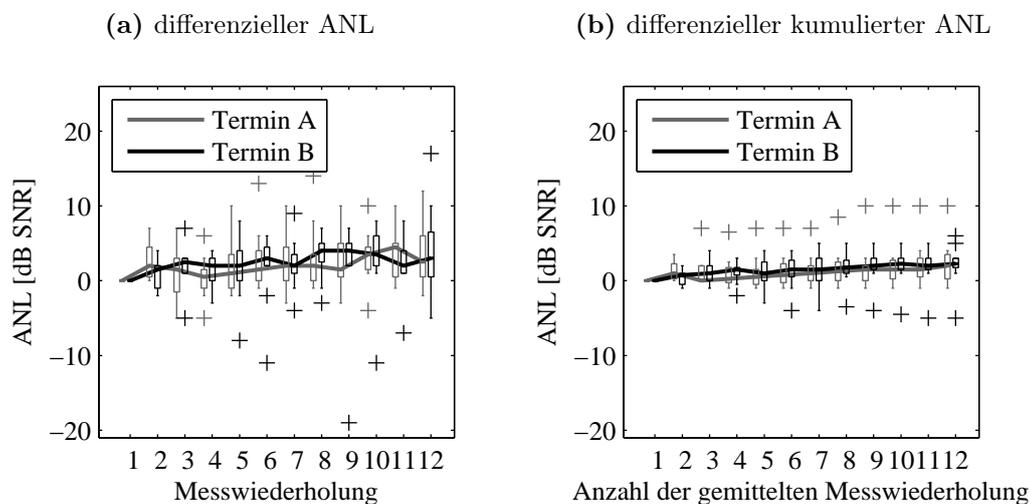


Abb. D.5: (a) ANL-Werte, gemessen bei bis zu zwölf Messwiederholungen. Die Werte sind durch eine Subtraktion des ersten ANL-Wertes jeder Versuchsperson von der interindividuellen Streuung befreit. (b) greift die Messwerte aus (a) auf und bildet den Median für jede Versuchsperson über die auf der Abszisse angegebene Anzahl der ersten Messungen.

D.4.5 Gegenüberstellung der JFC-Werte der Termine A und B

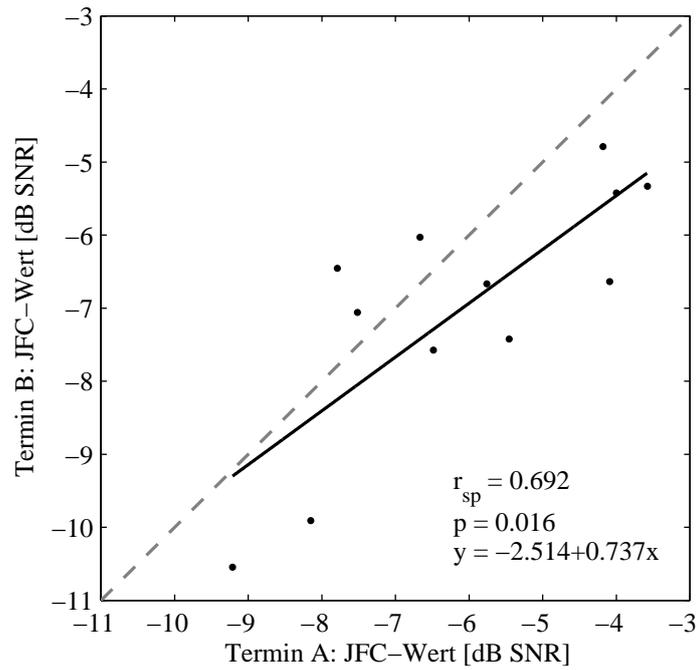


Abb. D.6: Gegenüberstellung der eingestellten JFC-Werte, gemessen in den Terminen A und B.

D.4.6 Inter- und intraindividuelle Streuung der JFC-Werte

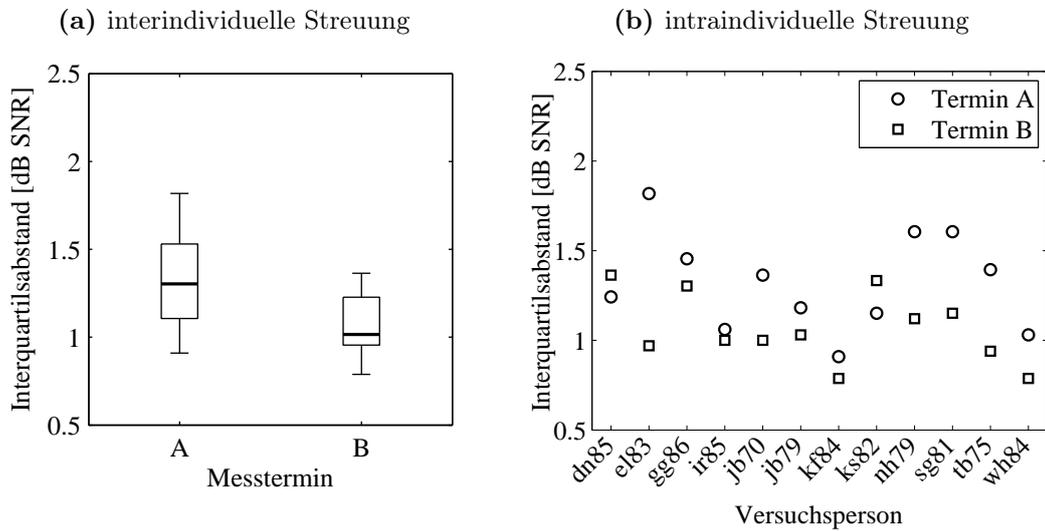


Abb. D.7: (a) interindividuelle Streuung und (b) die intraindividuelle Streuung der JFC-Werte, gemessen in den Terminen A und B.

D.4.7 Signifikanzprüfung der kumulierten JFC-Werte

Tab. D.5: Signifikante Unterschiede nach einem Friedman-Test zwischen den kumulierten JFC-Werten.

| Verglichene kumulierte Anzahl | $\chi^2_{(1;24;0,05)}$ | Irrtumswahr- scheinlichkeit p |
|----------------------------------|------------------------|----------------------------------|
| 1 und 5 | 4,226 | 0,04 |
| 1 und 6 | 4,19 | 0,042 |
| 1 und 7 | 4,668 | 0,031 |
| 1 und 8 | 4,589 | 0,032 |
| 1 und 9 | 4,154 | 0,042 |
| 1 und 10 | 4,589 | 0,032 |
| 1 und 11 | 4,668 | 0,031 |
| 1 und 12 | 4,084 | 0,043 |
| 2 und 12 | 4,05 | 0,044 |
| 3 und 10 | 4,154 | 0,0415 |
| 3 und 12 | 4,154 | 0,042 |
| 4 und 8 | 4,19 | 0,041 |

D.4.8 Berücksichtigung der interindividuellen Streuung im kumulierten JFC-Wert

Um eine Betrachtung der JFC-Werte in den Abb. 3.8 und 3.9 (s. Kap. 3.7) unabhängig von der interindividuellen Streuung der Ergebnisse zu ermöglichen, wurde der erste ermittelte JFC-Wert von den folgenden JFC-Werten jeder Versuchsperson subtrahiert. Die Ergebnisse sind in Abb. D.8a in Boxplots zusammengefasst. Abb. D.8b zeigt die kumulierten JFC-Werte der Abb. D.8a. Auch hier wurde der Median über die auf der Abszisse angegebene Anzahl der Messwerte ermittelt. Die Abbildungen zeigen, dass der Median der eingestellten JFC-Werte um ca. 1 dB SNR im Verlauf der Messung niedriger eingestellt wird als zu Beginn.

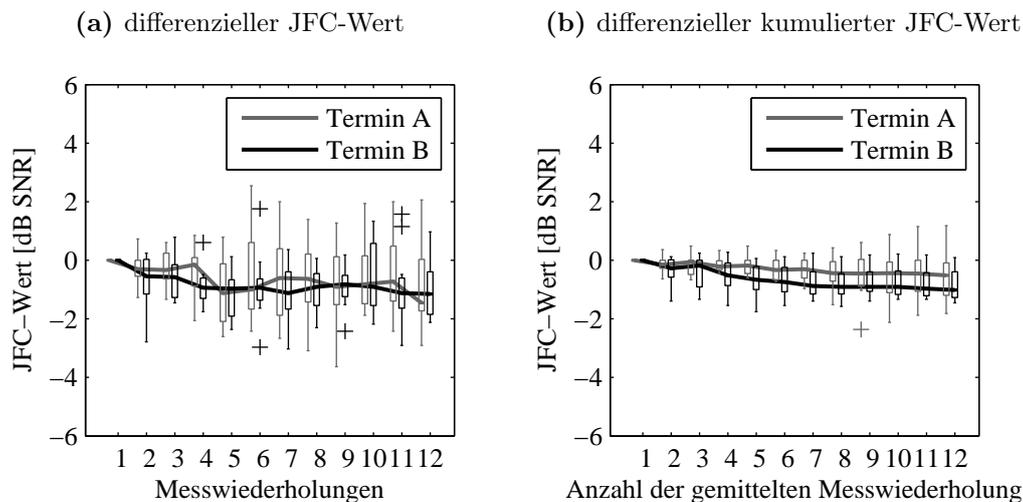


Abb. D.8: (a) JFC-Werte, gemessen bei bis zu zwölf Messwiederholungen in den Terminen A und B. Die Werte sind durch eine Subtraktion des ersten JFC-Wertes jeder Versuchsperson von der interindividuellen Streuung befreit. (b) greift die Messwerte aus (a) auf und bildet den Median für jede Versuchsperson über die auf der Abszisse angegebene Anzahl der ersten Messungen.

D.4.9 Abhängigkeit von ANLT und JFCT

In Abb. D.9 sind die Ergebnisse beider Versuchspersonengruppen mit der Messreihenfolge ANLT-JFCT wurden im Boxplot D.9a aufgetragen. In schwarz dargestellt sind die Ergebnisse der Gruppe 1, die im Termin A die Messungen in der Reihenfolge ANLT-JFCT durchführte und in grau die Ergebnisse der Gruppe 2, die Messungen in dieser Reihenfolge im Termin B durchführten. Die ANL-Werte, ermittelt in der Reihenfolge JFCT-ANLT, sind in Abb. D.9b dargestellt. In schwarz sind in dieser Abbildung die Werte der Gruppe 2 dargestellt. Sie führten im Termin A die Messungen in der Reihenfolge JFCT-ANLT durch. In grau ist das Ergebnis der Gruppe 1 dargestellt, die im Termin B die Reihenfolge JFCT-ANLT verwendeten. In der Abb. D.9a weichen die Median-Werte deutlicher voneinander ab als in Abb. D.9b.

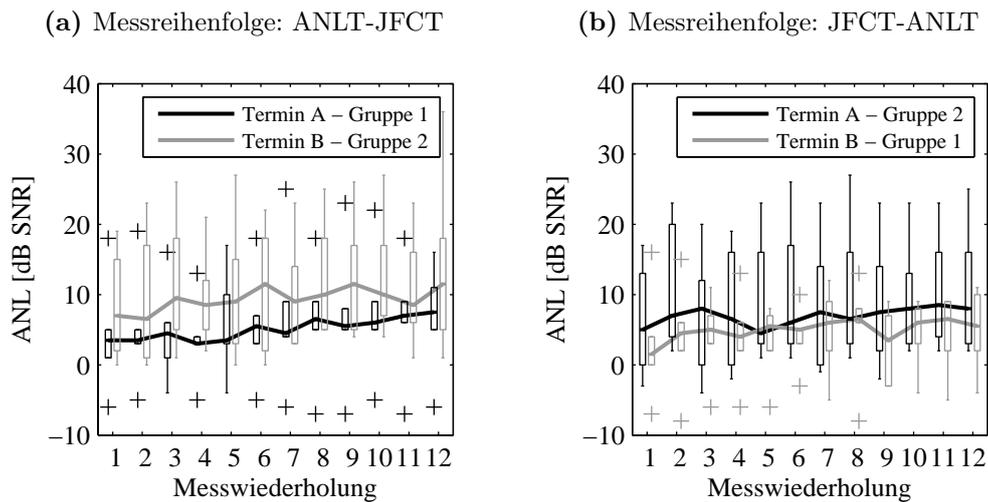


Abb. D.9: ANL-Werte der Versuchspersonen, gemessen in den Terminen A und B. (a) zeigt Werte, die in der Messreihenfolge ANLT-JFCT ermittelt wurden, (b) die Reihenfolge JFCT-ANLT.

Abb. D.10a und Abb. D.10b zeigen die JFC-Werte, ermittelt mit den Reihenfolgen ANLT-JFCT und JFCT-ANLT.

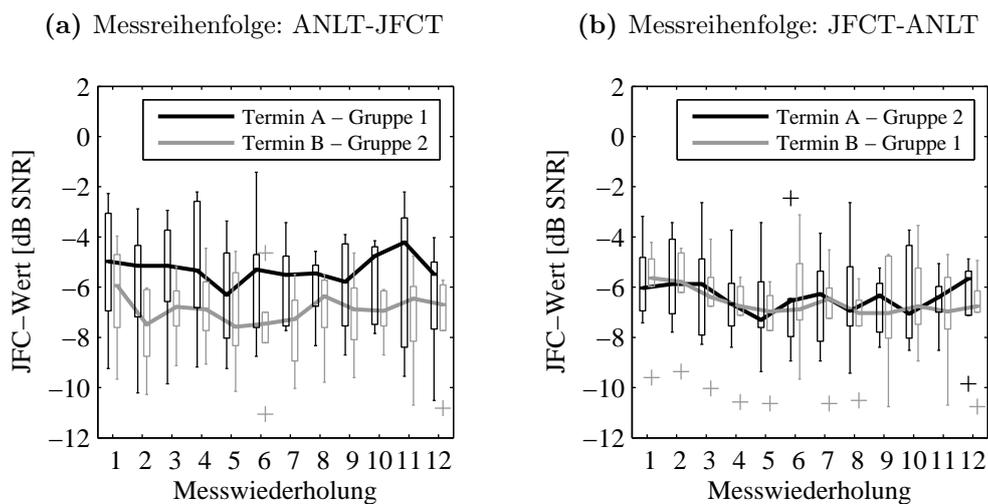


Abb. D.10: JFC-Werte der Versuchspersonen, gemessen in den Terminen A und B. (a) zeigt Werte, die in der Messreihenfolge ANLT-JFCT ermittelt wurden. In (b) ist die Reihenfolge JFCT-ANLT.

D.5 Experiment E: Einfluss der Methodik bei Wiederholungsmessungen auf den ANL

D.5.1 Gegenüberstellung der Messwerte im ANLT

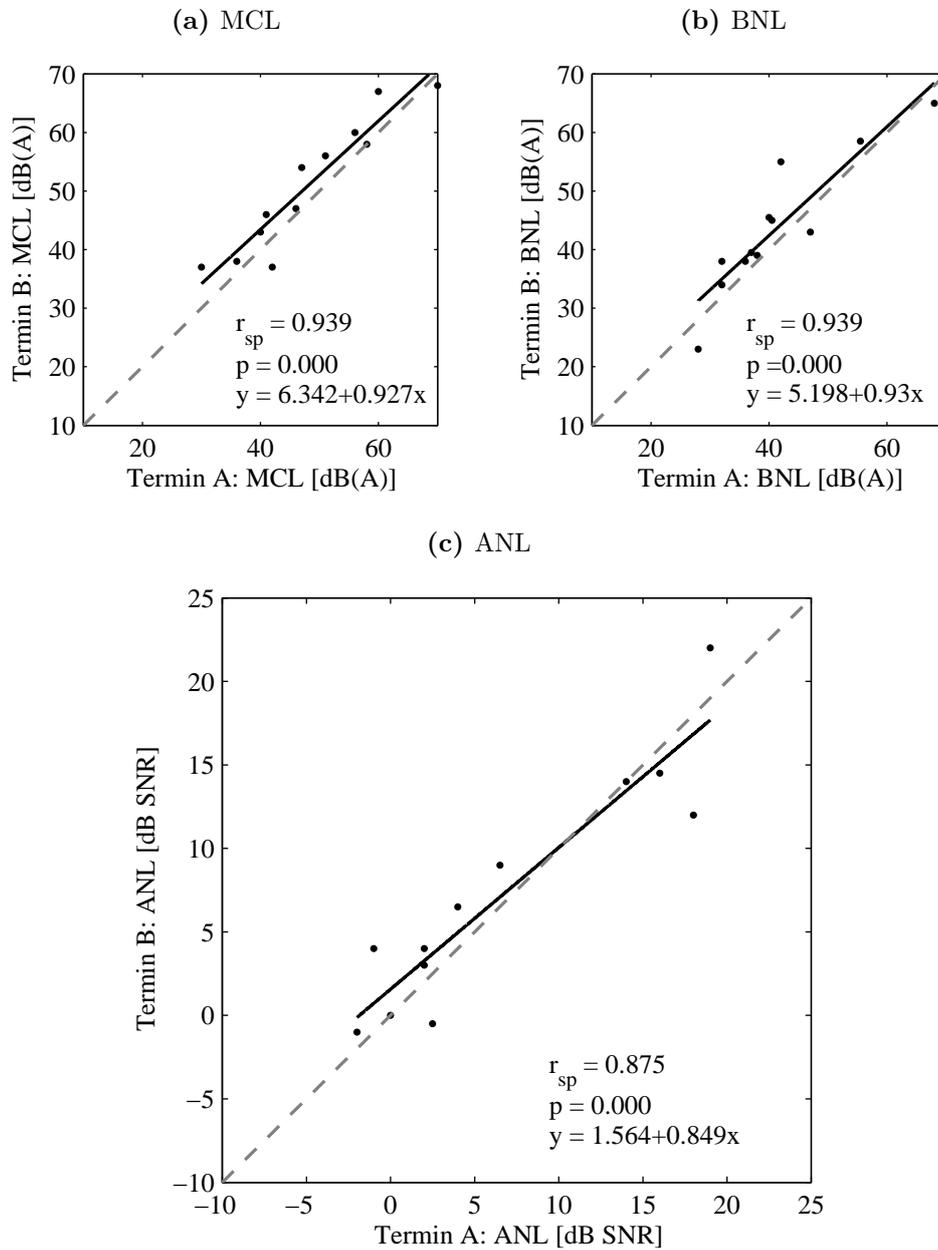


Abb. D.11: Gegenüberstellung des (a) MCL , (b) BNL und (c) ANL, gemessen in den Terminen A und B.

D.5.2 Inter- und intraindividuelle Streuung des ANL

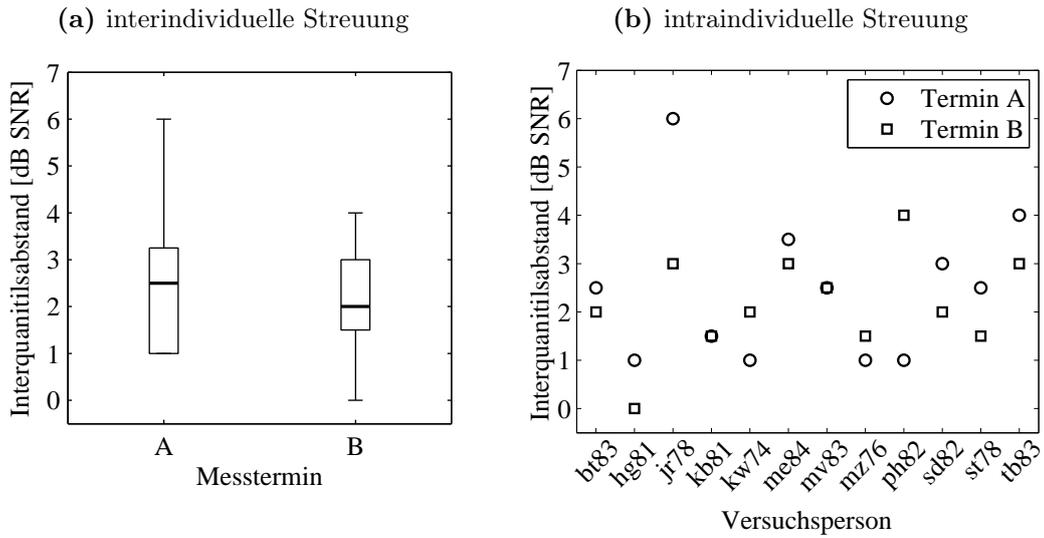


Abb. D.12: (a) zeigt die interindividuelle Streuung und (b) die intraindividuelle Streuung der ANL-Werte, gemessen in den Terminen A und B.

D.5.3 Signifikanzprüfung der kumulierten ANL-Werte

Tab. D.6: Signifikante Unterschiede nach einem Friedman-Test zwischen den kumulierten ANL-Werten ($\alpha = 0,05$).

| Verglichene kumulierte Anzahl | $\chi^2_{(1;24;0,05)}$ | Wahrscheinlichkeit p |
|-------------------------------|------------------------|------------------------|
| 2 und 5 | 4,459 | 0,035 |
| 2 und 6 | 4,750 | 0,029 |

D.5.4 Berücksichtigung der interindividuellen Streuung im kumulierten ANL

Um die Ergebnisse der zwölf Wiederholungsmessungen aus der Abb. 3.12c und 3.13 unabhängig von der interindividuellen Streuung zu betrachten, wurde von den Ergebnissen der erste ermittelte ANL jeder Versuchsperson subtrahiert. Die Ergebnisse sind in Abb. D.13a aufgeführt. Für Abb. D.13b wurden der Median aus der auf der Abszisse angegebenen Anzahl der ANL-Werte der Abb. D.13a ermittelt. Die dargestellten medianen Differenzen von ersten gemessenen ANL-Wert von maximal 1 dB SNR sind geringer als in Abb. D.5.

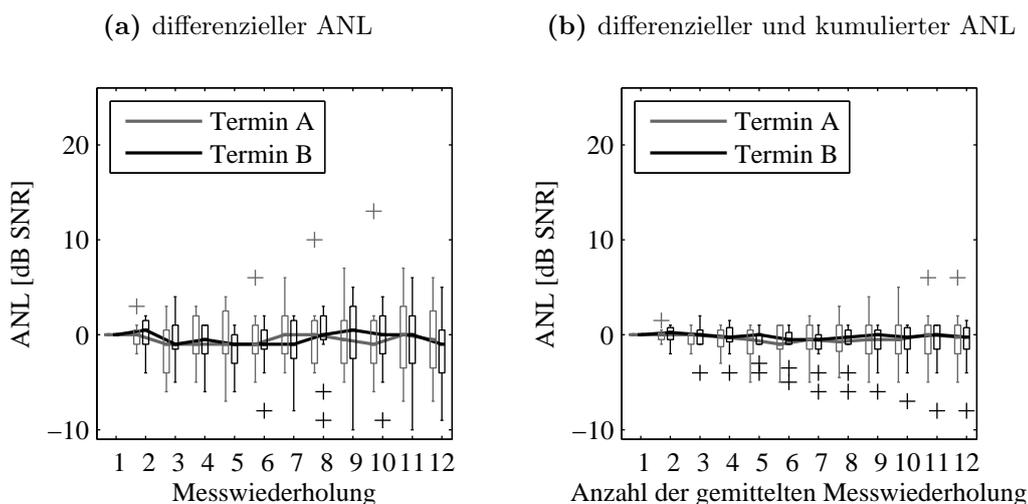


Abb. D.13: (a) zeigt die ANL-Werte, gemessen bei bis zu zwölf Messwiederholungen in den Terminen A und B. Die Werte sind durch Subtraktion des ersten ANL-Wertes jeder Versuchsperson von der interindividuellen Streuung befreit. (b) greift die Messwerte aus (a) auf und fasst den Median über die auf der Abszisse angegebene Anzahl der ersten Messungen jeder Versuchsperson zusammen.

D.6 Experiment G: Einfluss unterschiedlicher Hintergrundgeräusche

Tab. D.7: Ermittelte Irrtumswahrscheinlichkeiten p mit dem Wilcoxon Vorzeichen-Rang-Test bei der Prüfung der signifikanten Unterschiede zwischen den Medianen des ANL, ermittelt mit unterschiedlichen Hintergrundgeräuschen.

| Hintergrundgeräusche | OLSA | ANL | ISTS | Cafeteria | Maschine |
|----------------------|------|--------|--------|-----------|----------|
| OLSA | - | 0,0010 | 0,0596 | 0,4238 | 0,3296 |
| ANL | - | - | 0,0449 | 0,0015 | 0,0103 |
| ISTS | - | - | - | 0,0840 | 0,1938 |
| Cafeteria | - | - | - | - | 0,7188 |

D.7 Experiment H: Einfluss des Darbietungspegels bei der Übung des Sprachmaterials

Tab. D.8: Irrtumswahrscheinlichkeiten p bei der Prüfung der signifikanten Unterschiede zwischen den Messwerten des ANLT mit randomisiertem und konstantem Pegel in der Trainingsphase, ermittelt mit Mann-Whitney-U-Test und Kolmogorov-Smirnov-Test.

| Messwerte des ANLT | Mann-Whitney- U-Test | Kolmogorov- Smirnov-Test |
|-----------------------|-------------------------|-----------------------------|
| maximale Sprache | 0,748 | 0,786 |
| minimale Sprache | 0,977 | 0,991 |
| MCL | 0,751 | 0,786 |
| maximales Rauschen | 0,838 | 0,786 |
| minimales Rauschen | 0,767 | 0,991 |
| BNL | 0,452 | 0,433 |
| ANL | 0,181 | 0,433 |

Anhang E

Untersuchung von einkanaligen Störgeräuschreduktionen mit dem ANLT und anderen perzeptiven Maßen

E.1 ANL-Werte ermittelt mit Störgeräuschreduktion

Tab. E.1: ANL-Werte in dB SNR der Normalhörenden, gemessen mit unterschiedlichen Störgeräuschreduktionen

| Versuchspersonen | NoAlgo | Optimal | Real6dB | Real8dB |
|------------------|--------|---------|---------|---------|
| am77 | 17 | 12 | 17 | 17 |
| as78 | 11 | 7 | 6 | 10 |
| dv80 | 14 | 11 | 7 | 5 |
| er83 | 0 | -3 | 0 | 0 |
| fj84 | 8 | 5 | 7 | 6 |
| ih65 | 5 | 2 | 4 | 2 |
| nh79 | 19 | 17 | 19 | 17 |
| sm78 | 0 | 1 | 3 | 1 |
| so78 | 16 | 9 | 8 | 9 |
| ts80 | 17 | 10 | 12 | 12 |

Tab. E.2: ANL-Werte in dB SNR der Schwerhörigen, gemessen mit unterschiedlichen Störgeräuschreduktionen

| Versuchspersonen | NoAlgo | Optimal | Real6dB | Real8dB |
|------------------|--------|---------|---------|---------|
| af40 | 0 | -1 | 4 | 2 |
| ag54 | 5 | 0 | 6 | 7 |
| cb85 | 1 | -2 | 3 | 1 |
| ew42 | 4 | 0 | 5 | 2 |
| fk86 | -3 | -5 | -3 | -3 |
| kb93 | 9 | 7 | 7 | 6 |
| mf90 | -1 | -4 | -2 | -4 |
| mg64 | 7 | 3 | 7 | 6 |
| rs56 | 9 | 4 | 8 | 8 |
| wb40 | 0 | -6 | -2 | -3 |
| ww53 | 1 | -4 | -1 | -2 |

E.2 Signifikanzprüfung der perzeptiven Maße

Tab. E.3: Prüfung der signifikanten Unterschiede der perzeptiven Maße. Irrtumswahrscheinlichkeit p ermittelt mit einem Wilcoxon Vorzeichen-Rang-Test beim Vergleich der perzeptiven Maße, gemessen in Situationen ohne und mit der genannten Störgeräuschreduktion.

| perzeptives Maß | Störgeräuschreduktion | Normalhörende p | Schwerhörige p |
|-----------------|-----------------------|----------------------|---------------------|
| ANL | Optimal | 0,004 | 0,001 |
| | Real6dB | 0,078 | 0,914 |
| | Real8dB | 0,023 | 0,09 |
| JFC-Wert | Optimal | 0,002 | 0,001 |
| | Real6dB | 0,94 | 0,52 |
| | Real8dB | 0,711 | 0,7 |
| SRT | Optimal | 0,002 | 0,001 |
| | Real6dB | 0,797 | 0,898 |
| | Real8dB | 0,606 | 0,31 |
| PPDIS | Optimal | 0,067 | 0,278 |
| | Real6dB | 0,577 | 0,054 |
| | Real8dB | 0,353 | 0,353 |
| Präferenz | Optimal | 0,008 | 0,008 |
| | Real6dB | 0,031 | 0,008 |
| | Real8dB | 0,781 | 0,008 |

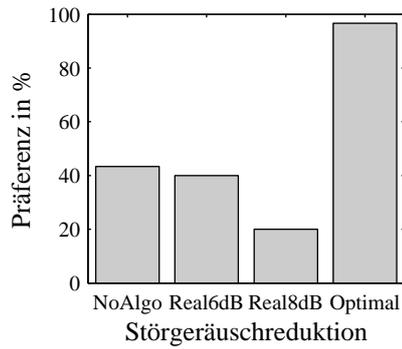
E.3 Paarvergleich

E.3.1 Ergebnisse der Paarvergleiche bei unterschiedlichen SNR-Werten

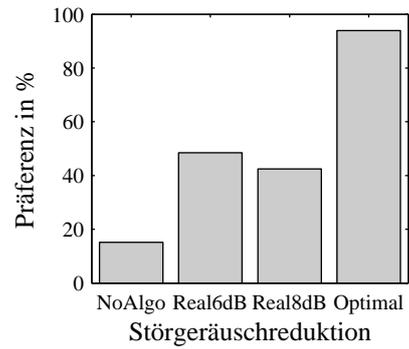
Zur Auswertung des Paarvergleiches wurden Häufigkeitsmatrizen für jeden gemessenen SNR erstellt. Dazu wurden die Ergebnisse jeder Versuchsperson in einer 4x4 Matrix dargestellt. Diese enthielt eine 1, wenn das Objekt der gewählten Zeile i dem Objekt der Spalte j vorgezogen wurde und 0 für den konträren Fall. Für ein Gesamturteil wurden die Matrizen aller Versuchspersonen zu Häufigkeitsmatrizen addiert. Jede Spalte wurde aufsummiert und durch die Anzahl der mit jeder Störgeräuschreduktion durchgeführten Vergleiche dividiert. Diese Ergebnisse sind als Präferenz in Prozent in Abb. E.1 für die SNR-Werte -3, 7,5, 12,5 und 19 dB SNR dargestellt. Sie zeigen, dass alle Versuchspersonen bei allen SNR-Werten die optimale Störgeräuschreduktion bevorzugen. Die Ergebnisse der Schwerhörigen, dargestellt in der rechten Spalte, zeigen, dass Situationen ohne Störgeräuschreduktion meist zurückgewiesen werden. Normalhörende Versuchspersonen lehnen bei den niedrigen SNR-Werten -3 dB SNR und 7,5 dB SNR die Störgeräuschreduktion Real8dB ab. Bei hohen SNR-Werten 12 dB SNR und 19 dB SNR werden Situationen ohne Störgeräuschreduktion abgewiesen. Die Konsistenz und die Akkordanz der Paarvergleiche sind in Anhang E.3.2 dargestellt.

Außerdem wurden im Paarvergleich die Signale bei einem SNR dargeboten, der den ANL-Werten, gemessen mit Störgeräuschreduktion, entsprach. In Abb. E.2 sind die Ergebnisse dieser Paarvergleiche dargestellt. In allen wird die Situation mit dem optimalen Algorithmus bevorzugt. Die schwerhörigen Versuchspersonen lehnen in der Regel die Situation ohne Störgeräuschreduktion ab. In Abb. E.2a zeigen Normalhörende ein ähnliches Verhalten. Die Signale wurden mit einem SNR dargeboten, der den ANL-Werten, ermittelt ohne Störgeräuschreduktion, entsprach. In allen anderen Paarvergleichen lehnen die Normalhörenden die Störgeräuschreduktion Real8dB in der Regel ab.

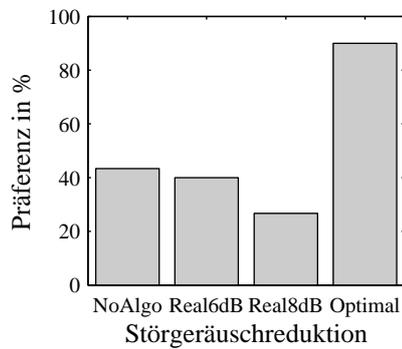
(a) SNR: -3 dB SNR, Normalhörende



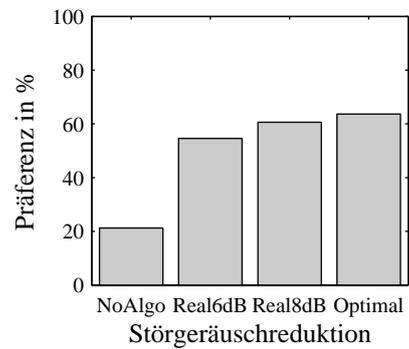
(b) SNR: -3 dB SNR, Schwerhörige



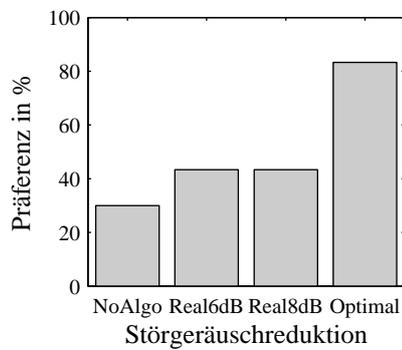
(c) SNR: 7,5 dB SNR, Normalhörende



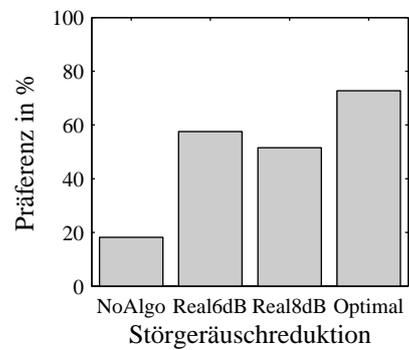
(d) SNR: 7,5 dB SNR, Schwerhörige



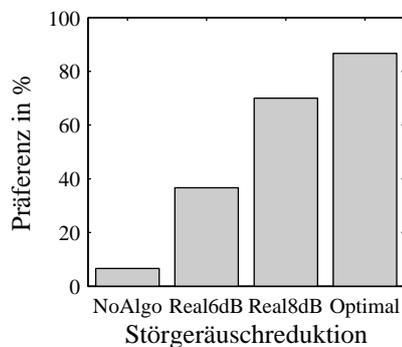
(e) SNR: 12,5 dB SNR, Normalhörende



(f) SNR: 12,5 dB SNR, Schwerhörige



(g) SNR: 19 dB SNR, Normalhörende



(h) SNR: 19 dB SNR, Schwerhörige

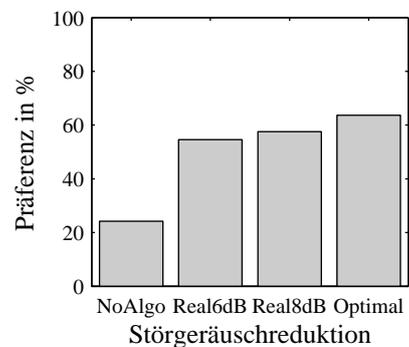
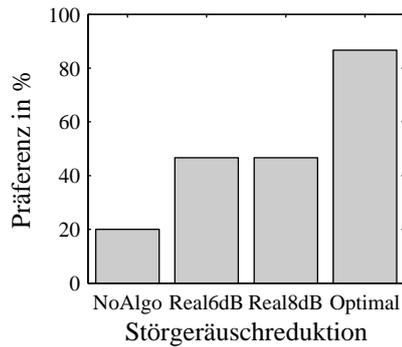
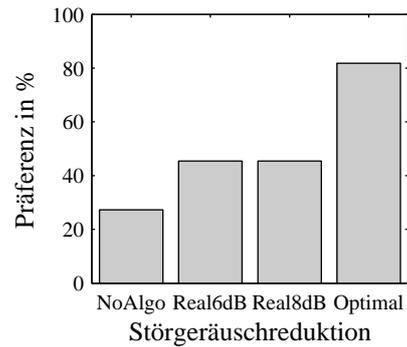


Abb. E.1: Ergebnisse des Paarvergleichs bei den SNR-Werten -3, 7,5, 12,5, 19 dB SNR für Normalhörende und Schwerhörige mit unterschiedlichen Störgeräuschreduktionen.

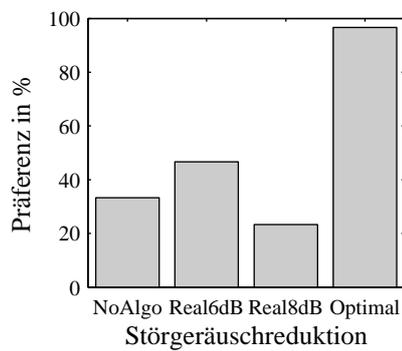
(a) SNR: ANL NoAlgo, Normalhörende



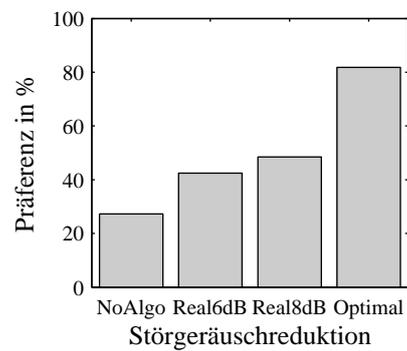
(b) SNR: ANL NoAlgo, Schwerhörige



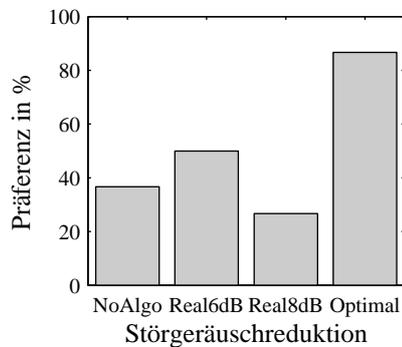
(c) SNR: ANL Real6dB, Normalhörende



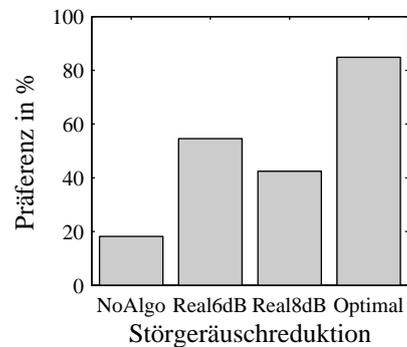
(d) SNR: ANL Real6dB, Schwerhörige



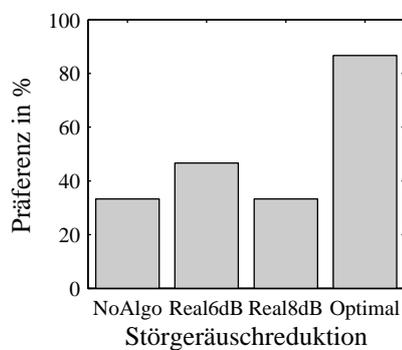
(e) SNR: ANL Real8dB, Normalhörende



(f) SNR: ANL Real8dB, Schwerhörige



(g) SNR: ANL Optimal, Normalhörende



(h) SNR: ANL Optimal, Schwerhörige

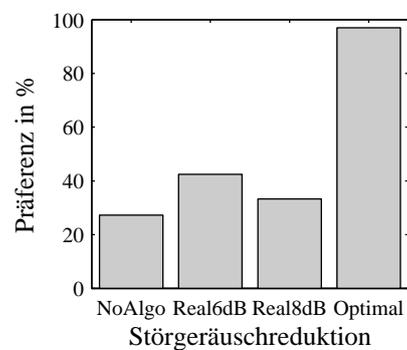


Abb. E.2: Ergebnisse des Paarvergleichs bei den SNRs ANL NoAlgo, ANL Real6dB, ANL Real8dB, ANL Optimal für die Normalhörenden und Schwerhörigen mit unterschiedlichen Störgeräuschreduktionen.

E.3.2 Prüfung der Konsistenz und Akkordanz im Paarvergleich

Die Tab. E.4 und E.5 zeigen den Median der Konsistenz jeder Versuchsperson in den Paarvergleichen. Für jeden Paarvergleich wurde eine Prüfung der Urteilkonsistenz durchgeführt (Bortz *et al.*, 2000). Ermittelt wurden die Irrtumswahrscheinlichkeiten p für eine zufällige Beurteilung zwischen 0,375 und 1. Alle Ergebnisse zeigen damit, dass eine signifikante Inkonsistenz bei einem Niveau von $\alpha = 0,01$ nicht nachgewiesen werden konnte.

Tab. E.4: Median der Konsistenz im Paarvergleich der normalhörenden Versuchspersonen

| Versuchsperson | Konsistenz |
|----------------|------------|
| er83 | 1 |
| sm78 | 1 |
| nh79 | 1 |
| dv80 | 1 |
| as78 | 1 |
| fj84 | 1 |
| ts80 | 1 |
| so78 | 1 |
| am77 | 1 |
| ih65 | 1 |

Tab. E.5: Median der Konsistenz im Paarvergleich der schwerhörigen Versuchspersonen

| Versuchsperson | Konsistenz |
|----------------|------------|
| kb93 | 0,75 |
| ww53 | 0,75 |
| af40 | 0,75 |
| cb85 | 0,75 |
| mf90 | 0,75 |
| fk86 | 1 |
| mg60 | 1 |
| wb40 | 0,75 |
| ew42 | 1 |
| ag54 | 0,5 |
| rs56 | 1 |

Die Tab. E.6 und E.7 zeigen die Akkordanz der Paarvergleiche bei gleichem SNR und die Ergebnisse einer Signifikanzprüfung. Alle Ergebnisse sind überzufällig identisch. Das Akkordanzmaß A ist aber z. T. sehr niedrig.

Tab. E.6: Akkordanz im Paarvergleich der normalhörenden Versuchspersonen

| SNR | A | χ^2 | Kritischer Wert | α |
|----------------|-------|----------|-----------------|----------|
| -3 | 0,452 | 38,438 | 19,187 | 0,99 |
| 12,5 | 0,178 | 19,938 | 19,187 | 0,99 |
| 19 | 0,593 | 47,938 | 19,187 | 0,99 |
| 7,5 | 0,281 | 26,938 | 19,187 | 0,99 |
| ANL6dB | 0,459 | 38,938 | 19,187 | 0,99 |
| ANL8dB | 0,237 | 23,938 | 19,187 | 0,99 |
| ANLNoAlgo | 0,289 | 27,438 | 19,187 | 0,99 |
| ANLWiener | 0,230 | 23,438 | 19,187 | 0,99 |
| Gesamtergebnis | 0,316 | 159,771 | 15,494 | 0,99 |

Tab. E.7: Akkordanz im Paarvergleich der schwerhörigen Versuchspersonen

| SNR | A | χ^2 | Kritischer Wert | α |
|----------------|-------|----------|-----------------|----------|
| -3 | 0,455 | 41,037 | 18,717 | 0,99 |
| 12,5 | 0,188 | 21,481 | 18,717 | 0,99 |
| 19 | 0,091 | 14,370 | 14,283 | 0,95 |
| 7,5 | 0,152 | 18,815 | 18,717 | 0,99 |
| ANL6dB | 0,176 | 20,593 | 18,717 | 0,99 |
| ANL8dB | 0,321 | 31,259 | 18,717 | 0,99 |
| ANLNoAlgo | 0,176 | 20,593 | 18,717 | 0,99 |
| ANLWiener | 0,406 | 37,481 | 18,717 | 0,99 |
| Gesamtergebnis | 0,243 | 136,071 | 15,456 | 0,99 |

E.4 Gegenüberstellung der perzeptiven Maße und der tatsächlichen SNR-Verbesserung

Abb. 4.3 zeigt, dass der Nutzen der Störgeräuschreduktionen abhängig vom SNR am Eingang ist. Deshalb wurde in Abb. E.3 der ΔANL in Abhängigkeit von diesem SNR-Wert aufgetragen. Die Grafen auf der linken Seite zeigen die Ergebnisse der Normalhörenden, die rechten Grafen die Ergebnisse der Schwerhörigen für die unterschiedlichen Störgeräuschreduktionen. Die Sterne markieren ΔANL -Werte. Dafür wurden auf der Ordinate der ΔANL und auf der Abszisse die dazugehörigen ANL-Werte, gemessen mit der jeweiligen Störgeräuschreduktion, aufgetragen. Diese ANL-Werte entsprechen den anliegenden SNR-Werten am Eingang der Störgeräuschreduktion. Zum Vergleich mit den ΔANL -Werten zeigt die durchgezogene Linie die SNR-Verbesserung für die jeweilige Störgeräuschreduktion aus Abb. 4.3. Für sie sind auf der Abszisse die SNR-Werte am Eingang der Störgeräuschreduktion SNR_{IN} aufgetragen und die Ordinate zeigt dann die tatsächliche SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$. Die Ergebnisse zeigen, dass die ΔANL -Werte stark um die $\text{SNR}_{\text{OUT-IN}}$ -Werte streuen und die Tendenzen der Median-Werte der Abb. 4.5 bestätigen.

In Abb. E.4 ist, wie für die ΔANL -Werte, der ΔJFC in Abhängigkeit vom SNR-Wert am Eingang der Störgeräuschreduktion zusammen mit der SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ aufgetragen. Die Ergebnisse zeigen, dass die ΔJFC -Werte weniger um die $\text{SNR}_{\text{OUT-IN}}$ -

Werte streuen als die ΔANL -Werte in Abb. E.3. Die JFC-Werte zeigen für die realen Störgeräuschreduktionen, dass die Messung bei SNR-Werten durchgeführt wurde, bei denen keine SNR-Verbesserung erreicht werden kann und eine Verbesserung durch die optimale Störgeräuschreduktion eher überschätzt wird.

In Abb. E.5 ist ebenfalls der ΔSRT zusammen mit der SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ in Abhängigkeit vom SNR-Wert am Eingang der Störgeräuschreduktion aufgetragen. Die Ergebnisse zeigen, dass die ΔSRT -Werte um die $\text{SNR}_{\text{OUT-IN}}$ -Werte weniger streuen als die ΔANL -Werte in Abb. E.3, aber in einem ähnlichen Maß, wie die des JFCT in Abb. E.4. Die Darstellung der ΔSRT -Werte zeigt für die realen Störgeräuschreduktionen, dass auch diese Messungen bei SNR-Werten durchgeführt wurden, bei denen keine SNR-Verbesserung erreicht werden kann und die optimale Störgeräuschreduktion eher überschätzt wird.

Für die Abb. E.6 ist die Anzahl der Bevorzugen gegenüber NoAlgo bei den Paarvergleichen mit den festen SNR-Werten von -3, 7,5, 12,5 und 19 dB SNR ermittelt worden. Diese Werte wurden dann mit Hilfe der zusätzlich als durchgezogene Linie dargestellten $\text{SNR}_{\text{OUT-IN}}$ -Werte wie folgt normiert. Die maximal mögliche Anzahl der Bevorzugen entspricht dem maximalen $\text{SNR}_{\text{OUT-IN}}$ -Wert. Keine Bevorzugen erhielt den Wert 0. Abb. E.6 zeigt, dass die Ergebnisse der Paarvergleiche dem $\text{SNR}_{\text{OUT-IN}}$ -Verlauf nicht folgen können.

Tab. E.8 zeigt die Spearmanscher Korrelationskoeffizienten r_{sp} und die Irrtumswahrscheinlichkeiten p für die Zusammenhänge zwischen den perzeptiven Maßen und der tatsächlichen SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ durch die Störgeräuschreduktionen. Zur Berechnung der Korrelation wurde zu jedem eingestellten perzeptiven Maß (ΔANL , ΔJFC -Wert, ΔSRT , Präferenz gegenüber *NoAlgo*) der $\text{SNR}_{\text{OUT-IN}}$ -Wert bei gleichem SNR am Eingang der Störgeräuschreduktion ermittelt.

Tab. E.8: Spearmanscher Korrelationskoeffizient r_{sp} und Irrtumswahrscheinlichkeit p für den Zusammenhang zwischen ΔANL , ΔJFC , ΔSRT und der Anzahl an Bevorzugen gegenüber NoAlgo aus dem Paarvergleich und der tatsächlichen SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ durch die Störgeräuschreduktionen.

| perzeptives Maß | Störgeräuschreduktion | Normalhörende | | Schwerhörige | |
|--|-----------------------|-----------------|-------|-----------------|-------|
| | | r_{sp} | p | r_{sp} | p |
| ΔANL | Optimal | -0,182 | 0,615 | 0,046 | 0,892 |
| | Real6dB | 0,542 | 0,106 | 0,492 | 0,124 |
| | Real8dB | 0,367 | 0,297 | -0,380 | 0,249 |
| ΔJFC -Wert | Optimal | -0,085 | 0,815 | 0,340 | 0,306 |
| | Real6dB | 0,043 | 0,907 | 0,018 | 0,968 |
| | Real8dB | 0,067 | 0,854 | -0,291 | 0,386 |
| ΔSRT | Optimal | -0,440 | 0,203 | -0,096 | 0,780 |
| | Real6dB | -0,312 | 0,380 | 0,105 | 0,759 |
| | Real8dB | -0,505 | 0,137 | -0,416 | 0,203 |
| Anzahl der Bevorzugen gegenüber NoAlgo | Optimal | 0,316 | 1 | 0,949 | 0,167 |
| | Real6dB | -0,105 | 1 | -0,211 | 0,833 |
| | Real8dB | 0,200 | 0,917 | -0,894 | 0,333 |

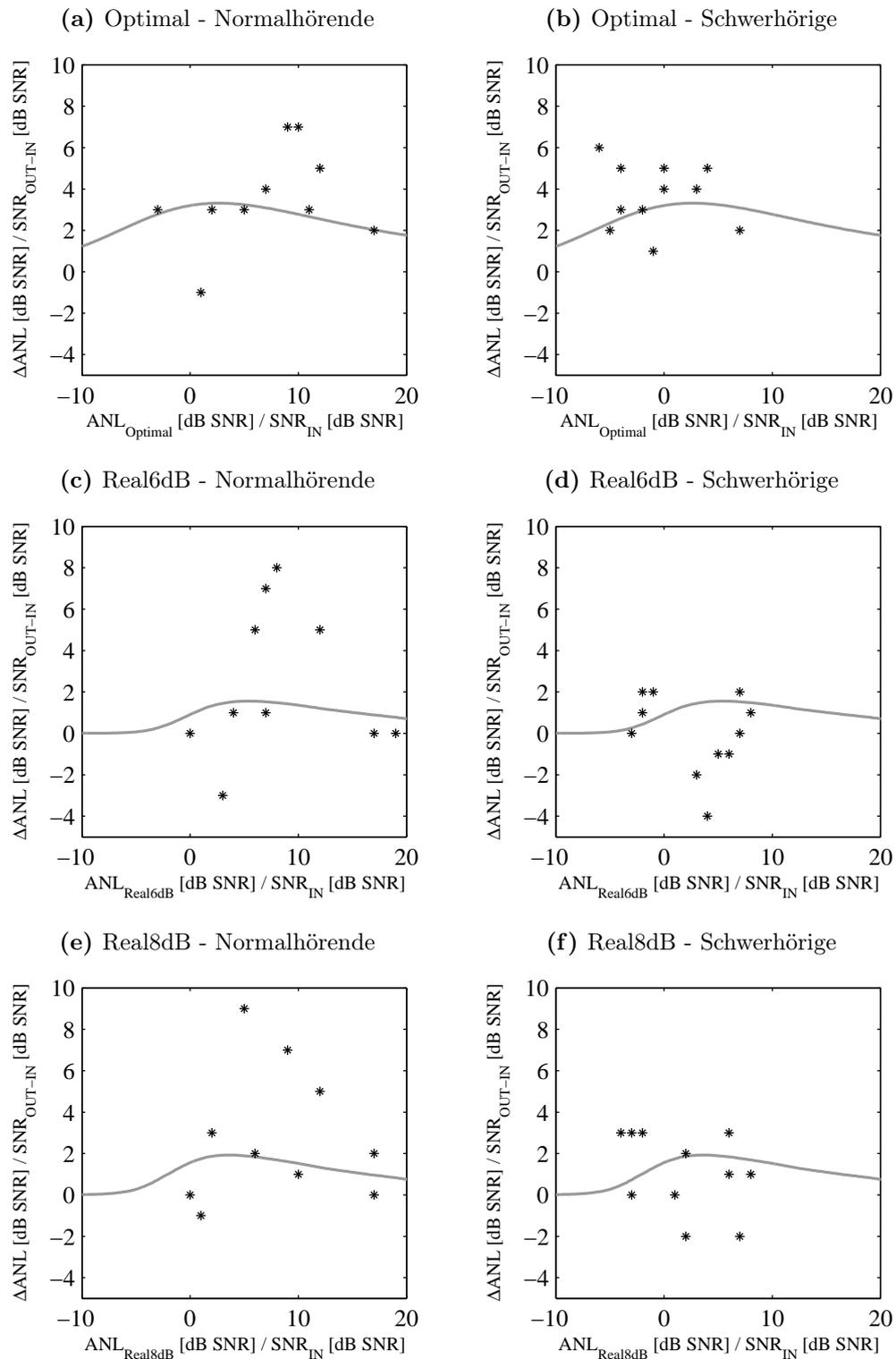


Abb. E.3: Vergleich der ΔANL -Werte mit der SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ in Abhängigkeit vom dargebotenen SNR. Dargestellt sind die ΔANL -Werte, die die Versuchspersonen für die verschiedenen Störgeräuschreduktionen eingestellt haben (Sterne). Die durchgezogene Linie zeigt die SNR-Verbesserung in Abhängigkeit vom SNR am Eingang der Störgeräuschreduktion aus Abb. 4.3.

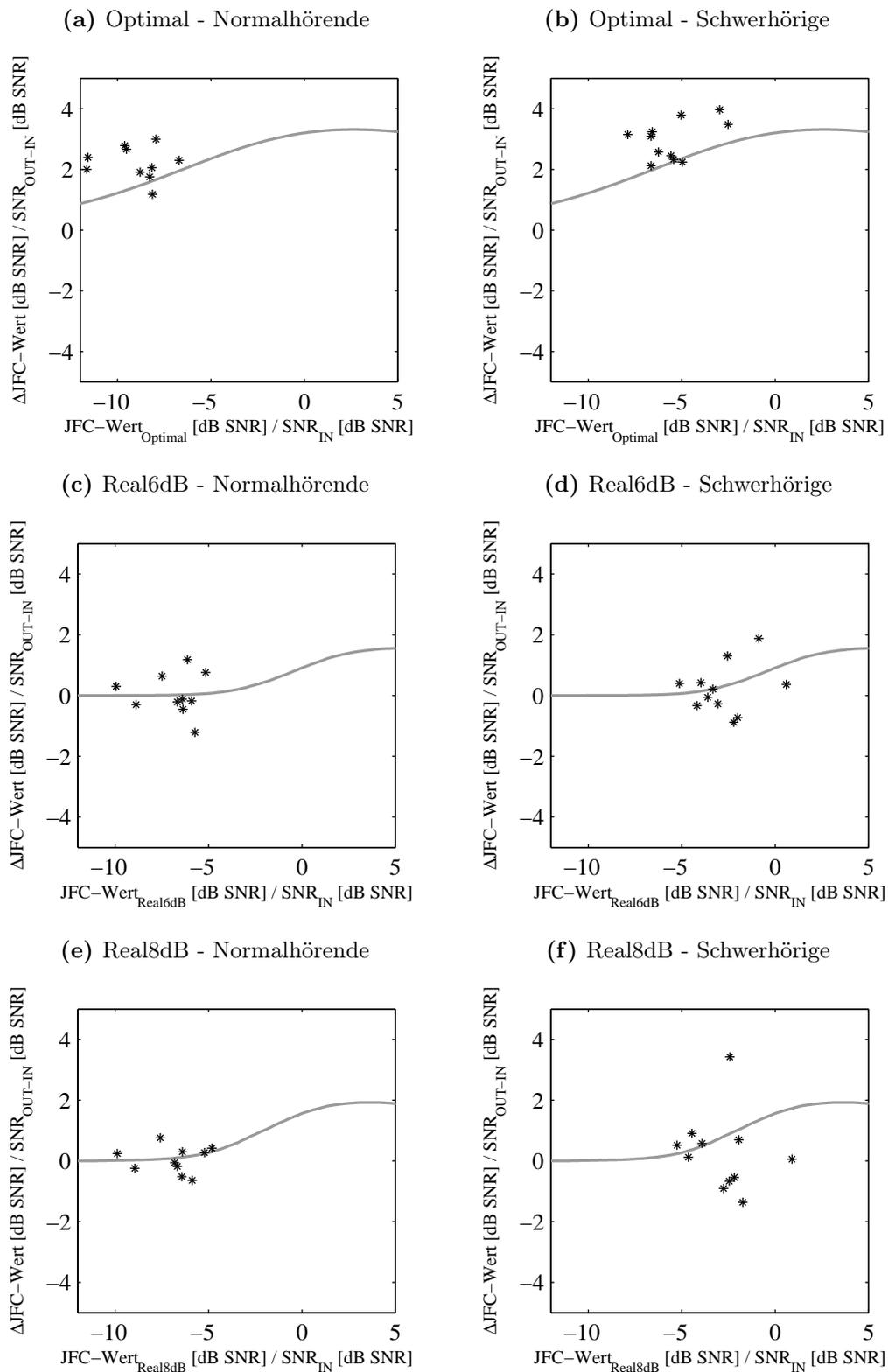


Abb. E.4: Vergleich der ΔJFC -Werte mit der SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ in Abhängigkeit vom dargebotenen SNR. Dargestellt sind die ΔJFC -Werte, die die Versuchspersonen für die verschiedenen Störgeräuschreduktionen eingestellt haben (Sterne). Die durchgezogene Linie zeigt die SNR-Verbesserung in Abhängigkeit vom SNR am Eingang der Störgeräuschreduktion aus Abb. 4.3.

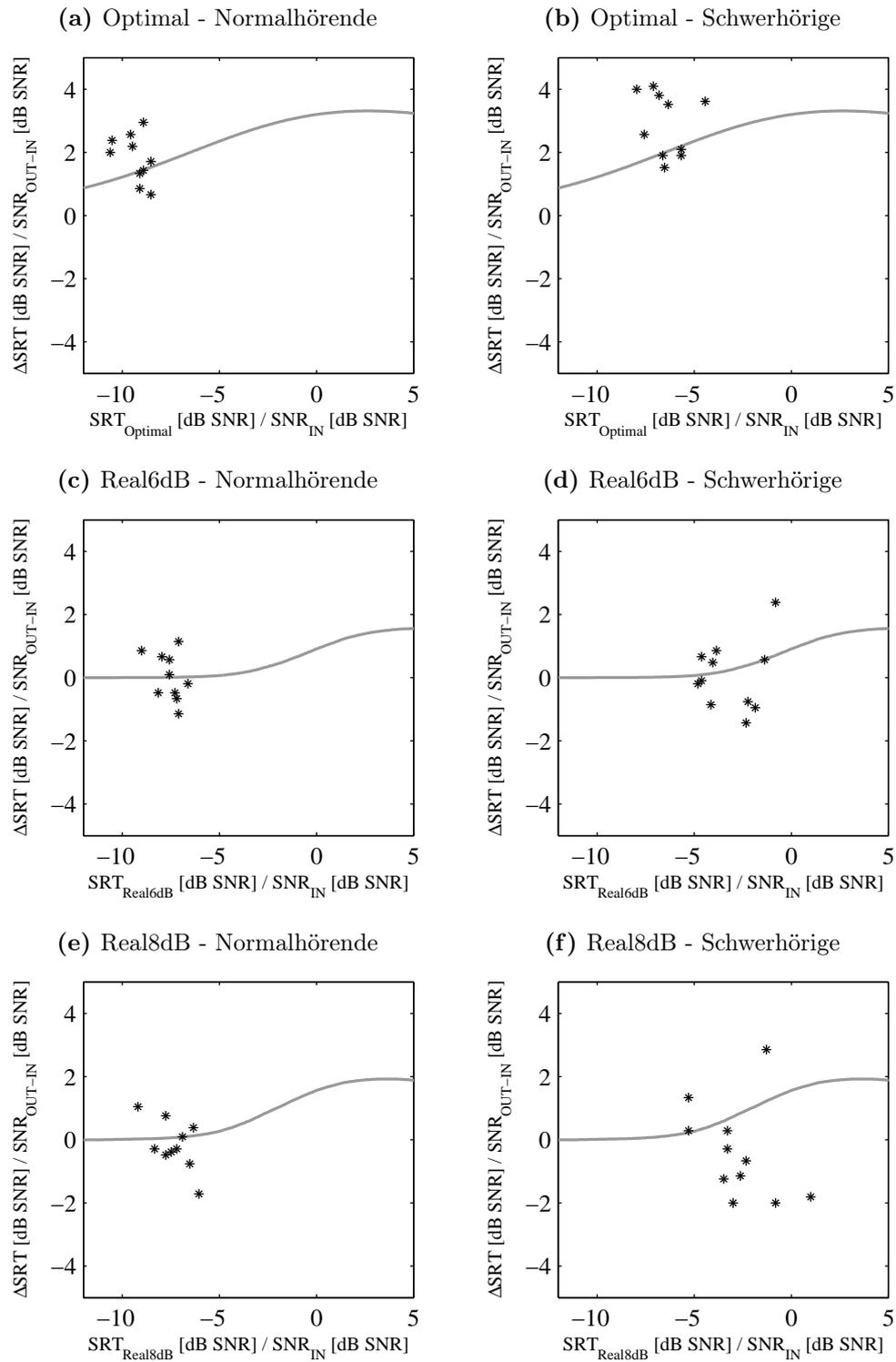


Abb. E.5: Vergleich der Δ SRT-Werte mit der SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ in Abhängigkeit vom dargebotenen SNR. Dargestellt sind die Δ SRT-Werte, die die Versuchspersonen für die verschiedenen Störgeräuschreduktionen eingestellt haben (Sterne). Die durchgezogene Linie zeigt die SNR-Verbesserung in Abhängigkeit vom SNR am Eingang der Störgeräuschreduktion aus Abb. 4.3.

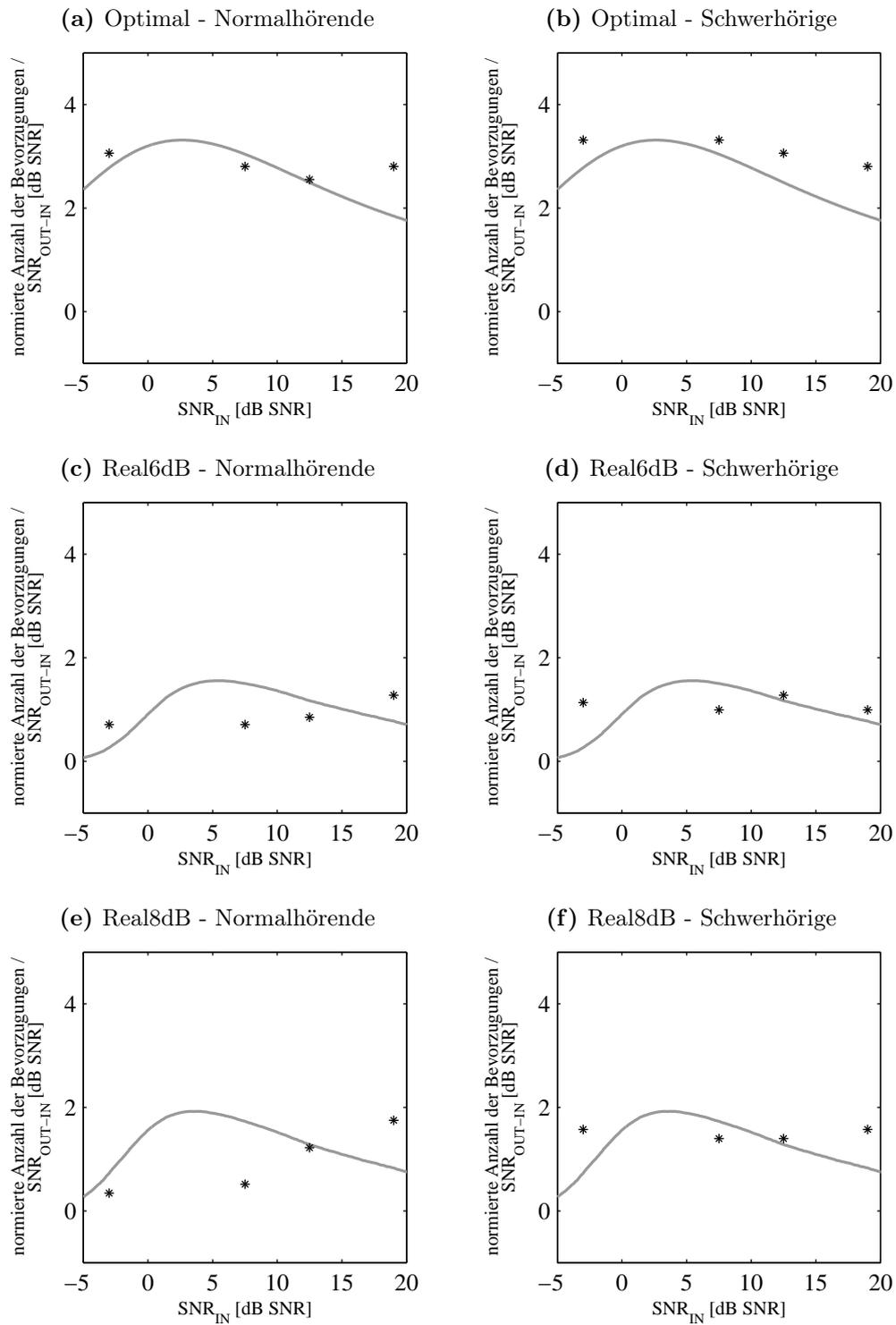
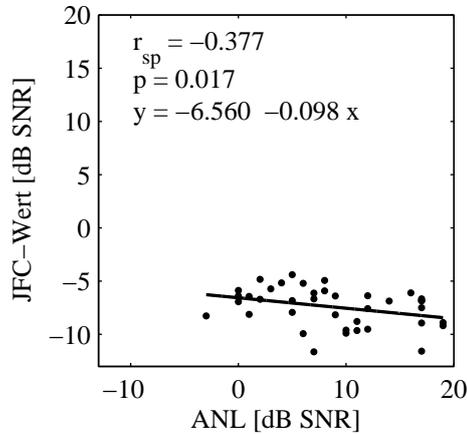


Abb. E.6: Vergleich der normierten Anzahl der Bevorzugungen gegenüber NoAlgo mit der SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ in Abhängigkeit vom dargebotenen SNR. Dargestellt ist die normierte Anzahl der Bevorzugungen bei festen SNR-Werten (Sterne). Die durchgezogene Linie zeigt die SNR-Verbesserung in Abhängigkeit vom SNR am Eingang der Störgeräuschreduktion aus Abb. 4.3.

E.5 Vergleich der ANL-Werte mit den anderen perzeptiven Maßen

(a) ANL - JFC-Wert (Normalhörende)



(b) ANL - JFC-Wert (Schwerhörige)

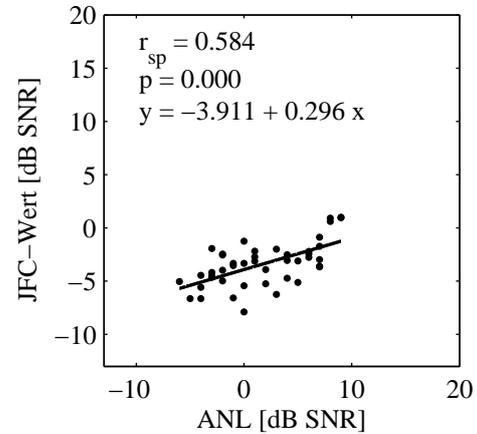
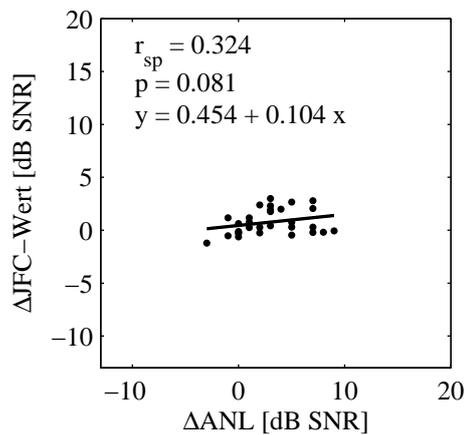
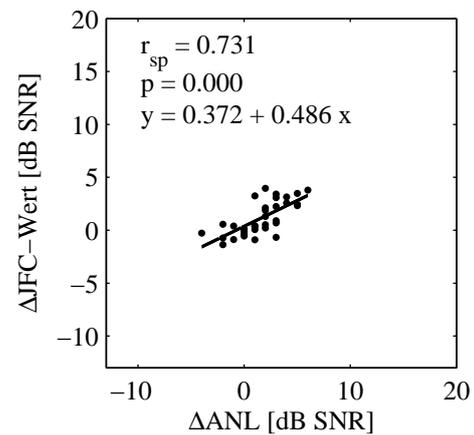
(c) Δ ANL - Δ JFC-Wert (Normalhörende)(d) Δ ANL - Δ JFC-Wert (Schwerhörige)

Abb. E.7: Vergleich der ermittelten ANL- und JFC-Werte. (a) und (b) stellen die ermittelten ANL- und JFC-Werte der Normalhörenden und der Schwerhörigen gegenüber. (c) und (d) zeigen die Δ ANL- und Δ JFC-Werte der Normalhörenden und Schwerhörigen.

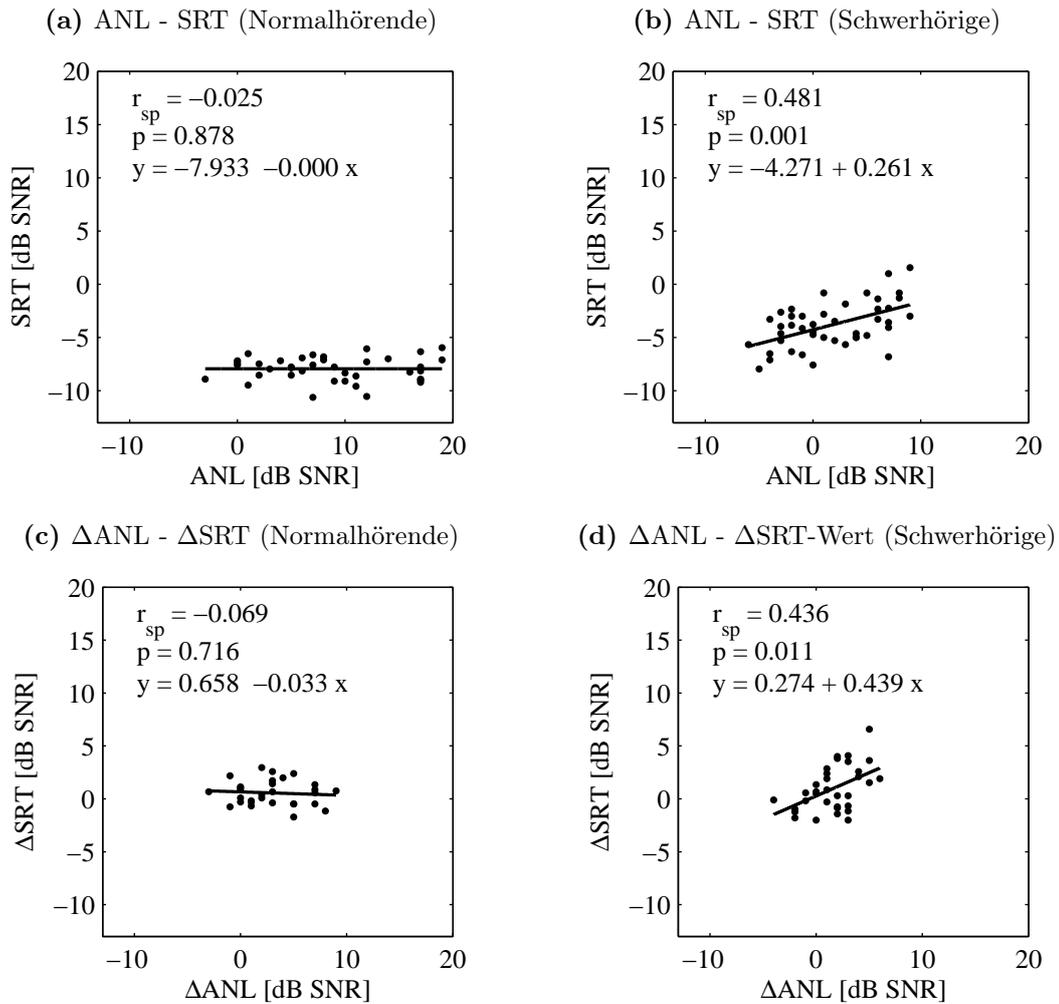


Abb. E.8: Vergleich der ermittelten ANL- und SRT-Werte. (a) und (b) stellen die ermittelten ANL- und SRT-Werte der Normalhörenden und der Schwerhörigen gegenüber. (c) und (d) zeigen die Δ ANL- und Δ SRT-Werte der Normalhörenden und Schwerhörigen.

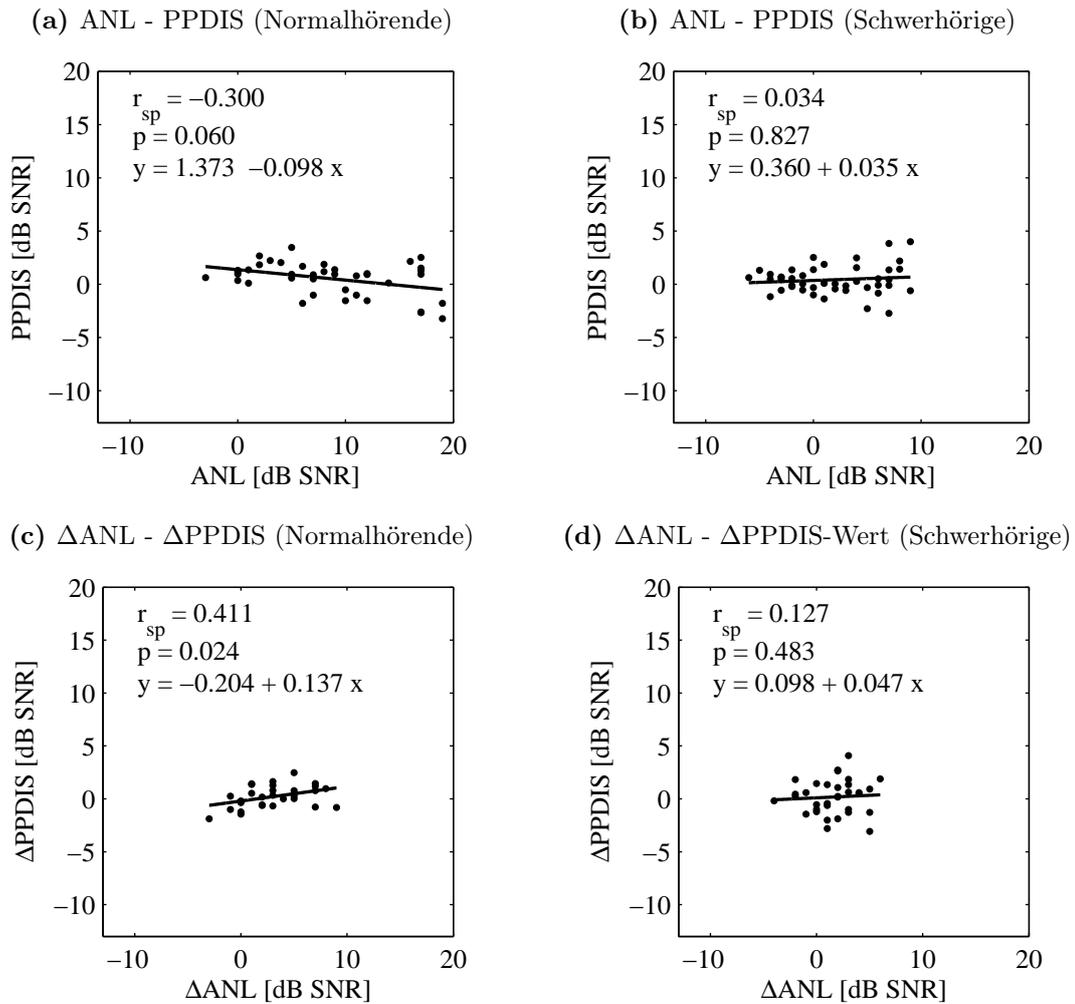
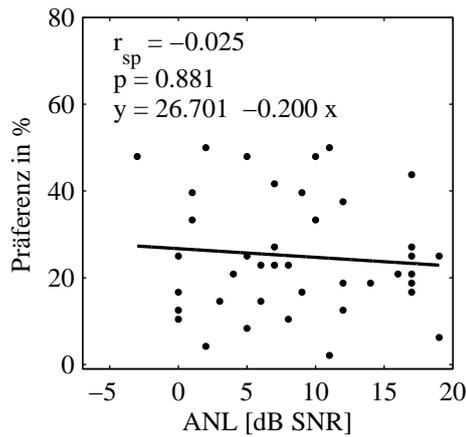


Abb. E.9: Vergleich der ermittelten ANL- und PPDIS-Werte. (a) und (b) stellen die ermittelten ANL- und PPDIS-Werte der Normalhörenden und der Schwerhörigen gegenüber. (c) und (d) zeigen die Δ ANL- und Δ PPDIS-Werte der Normalhörenden und Schwerhörigen.

(a) ANL - Präferenz (Normalhörende)



(b) ANL - Präferenz (Schwerhörige)

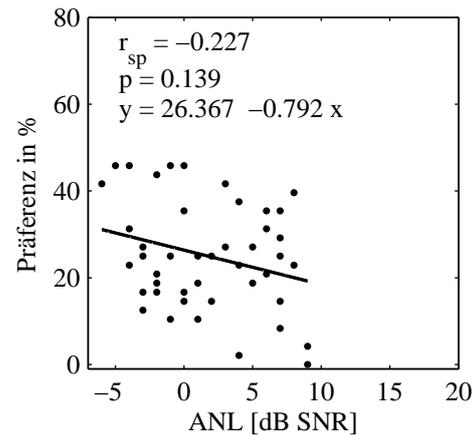
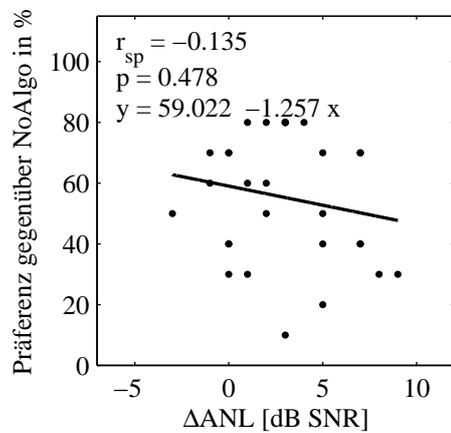
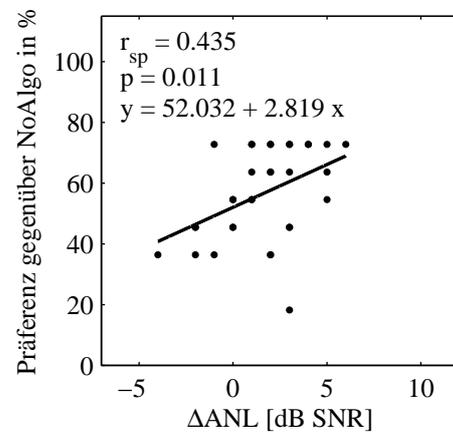

 (c) Δ ANL - Präferenz gegenüber NoAlgo (Normalhörende)

 (d) Δ ANL - Präferenz gegenüber NoAlgo (Schwerhörige)


Abb. E.10: Vergleich der ermittelten ANL-Werte und Werte des Paarvergleichs für die Ergebnisse der Normalhörenden und der Schwerhörigen. (a) und (b) stellen die ermittelten ANL-Werte den Präferenzen des Paarvergleiches gegenüber. (c) und (d) zeigen die Δ ANL und Präferenz gegenüber NoAlgo.

Anhang F

Untersuchung von einkanaligen Störgeräuschreduktionen mit dem ANLT und zwei technischen Maßen

Um die Ergebnisse des ANLT mit den ermittelten Δ SNR-Werten nach Hagerman und Olofsson (2004) zu vergleichen, wurden die Δ ANL-Werte (Sterne) und Δ SNR-Werte (durchgezogene Linie) in Abb. F.1 über dem Eingangs-SNR SNR_{IN} dargestellt. Die linke Spalte der Abbildungen zeigt die Δ ANL-Werte der Normalhörenden bei unterschiedlichen Störgeräuschreduktionen, die rechte Spalte die Ergebnisse der Schwerhörigen. Außerdem zeigt die unterbrochene Linie die tatsächliche SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$.

Für den Vergleich der perzeptiven Maße des ANLT der Normalhörenden und Schwerhörigen mit der MTF bei unterschiedlichen Störgeräuschreduktionen, wurden die Δ ANL-Werte (Sterne) mit der normierten und logarithmierten MTF (durchgezogene Linie) in Abb. F.2 dargestellt. Die linke Spalte der Abbildungen zeigt die Δ ANL-Werte der Normalhörenden bei unterschiedlichen Störgeräuschreduktionen, die rechte Spalte die Ergebnisse der Schwerhörigen. Außerdem zeigt die unterbrochene Linie die tatsächliche SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$.

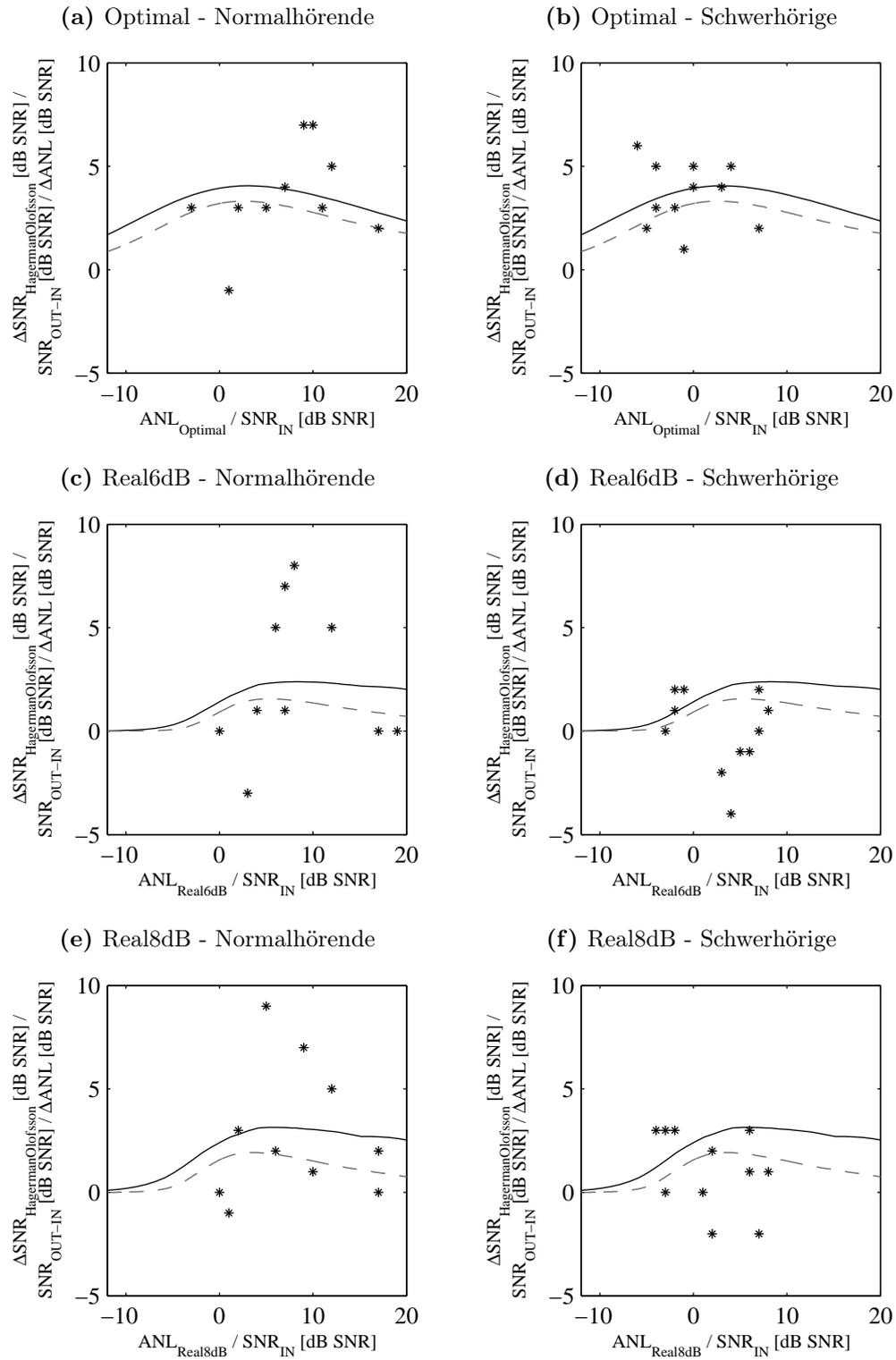


Abb. F.1: ΔSNR -Werte (durchgezogene Linie) und ΔANL -Werte (Sterne), aufgetragen über dem SNR am Eingang SNR_{IN} der Störgeräuschreduktionen. Die unterbrochene Linie zeigt die tatsächliche SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ jeder Störgeräuschreduktion aus Abb. 5.3.

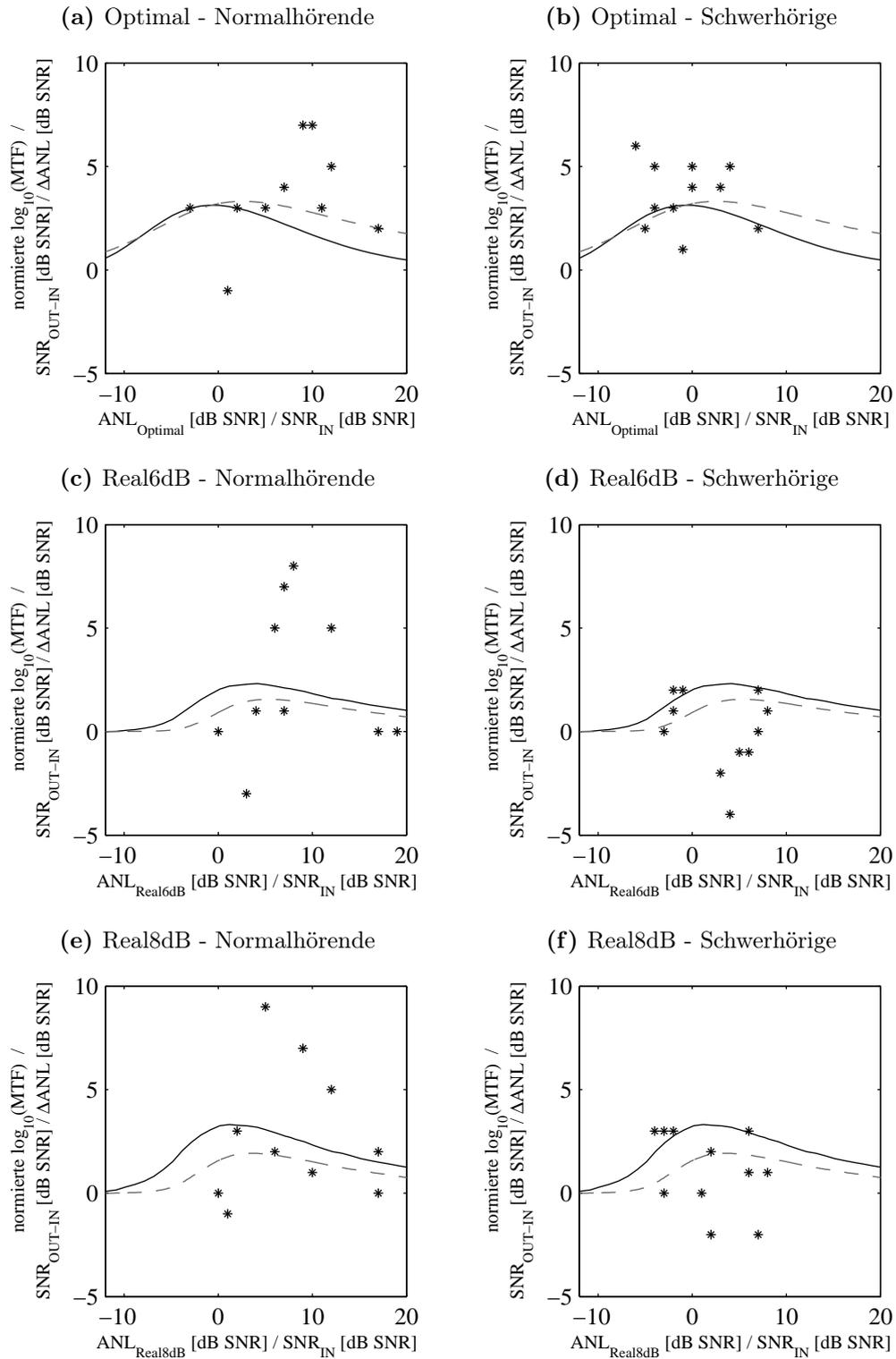


Abb. F.2: Normierte und logarithmierte MTF, aufgetragen über den SNR-Werten am Eingang der Störgeräuschreduktion (durchgezogene Linie) zusammen mit den ΔANL -Werten der Normalhörenden und Schwerhörigen (Sterne). Die unterbrochene Linie zeigt die tatsächliche SNR-Verbesserung $\text{SNR}_{\text{OUT-IN}}$ jeder Störgeräuschreduktion aus Abb. 5.3.

Literaturverzeichnis

- Alcántara, J. I., B. C. J. Moore, V. Kühnel und S. Launer (2003). Evaluation of the noise reduction system in a commercial digital hearing aid. *International Journal of Audiology*, **42**:34–42.
- ANSI S 3.5 - 1997 (1997). *American national standard methods for the calculation of the articulation index*. New York.
- Appiah, M. Y., R. Makrickaite, M. Gusaite und S. Munagala (2005). *Robust Voice Activity Detection and Noise Reduction Mechanism Using Higher-Order Statistics*. Technical Report: Aalborg University, Institute of Electronic Systems, Department of Communication Technology.
- Bentler, R. und L. Chiou (2006). Digital Noise Reduction: An Overview. *Trends in Amplification*, **10**(2):67–82.
- Bilger, R., J. Neutzel, M. Rabinowitz und C. Rzeczkowski (1984). Standardizing of a test of speech perception in noise. *Journal of Speech and Hearing Research*, **27**:32–48.
- Borg, E., M. Wilson und E. Samuelsson (1998). Towards an ecological audiology: stereophonic listening chamber and acoustic environmental tests. *Scandinavian Audiology*, **27**(4):195–206.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler* (5. Aufl.). Berlin: Springer Verlag.
- Bortz, J. und N. Döring (2005). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (3. überarbeitete Aufl.). Heidelberg: Springer Medizin Verlag.
- Bortz, J., G. A. Linert und K. Bohnke (2000). *Verteilungsfreie Methoden der Biostatistik* (2. Aufl.). Berlin: Springer Verlag.
- Boymans, M. und W. A. Dreschler (2000). Field Trials Using a Digital Hearing Aid with Active Noise Reduction and Dual-Microphone Directionality. *Audiology*, **39**:260–268.
- Boymans, M., W. A. Dreschler, P. Schoneveld und H. Verschuure (1999). Clinical Evaluation of a Full-digital In-the-ear Hearing Instrument. *Audiology*, **38**(2):99–108.
- Bray, V. und M. Nilsson (2001). Additive Snr Benefits of Signal Processing Features in a Directional DSP Aid. *The Hearing Review*, **8**(12):48–51,62.
- Bray, V., S. Sandridge, C. Newman und S. Kornhass (2002). Clinical research findings confirm benefit of advanced signal processing. *Audiology Online*.

- Campione, E. und J. Veronis (2002). A large-scale multilingual study of silent pause duration. In *Speech Prosody 2002*, S. 199–202. Aix-en-Provence, France.
- Cappé, O. (1994). Elimination of the Musical Noise Phenomenon with the Ephraim and Malah Noise Suppressor. *IEEE Transactions on Speech and Audio Processing*, **2(2)**:345–349.
- Cohen, I. und B. Berdugo (2002). Noise Estimation by Minima Controlled Recursive Averaging for Robust Speech Enhancement. *IEEE Signal Processing Letters*, **9(1)**:12–15.
- Crowley, H. J. und A. K. Nabelek (1996). Estimation of client-assessed hearing aid performance based upon unaided variables. *Journal of Speech and Hearing Research*, **39**:19–27.
- DIN EN 60645-2 (1997). *Audiometer - Teil 2: Geräte für die Sprachaudiometrie*.
- DIN EN ISO 389-8 (2004). *Akustik - Standard-Bezugspegel für die Kalibrierung audiometrischer Geräte - Teil 8: Äquivalente Bezugs-Schwellenschalldruckpegel für reine Töne und circumaurale Kopfhörer*.
- Drullmann, R., J. Festen und R. Plomp (1994). Effect of reducing slow temporal modulations on speech reception. *Journal of the Acoustical Society of America*, **95(5)**:2670–2680.
- Eberling, C., C. Ludvigsen und G. Keidser (1993). The Design and Testing of a Noise Reduction Algorithm Based on Spectral Subtraction. *Scandinavian Audiology*, **Suppl. 38**:39–49.
- Edwards, A. (1957). *Techniques of Attitude Scale Construction*. New York: Appelton-Century-Crofts.
- Ephraim, Y. und D. Malah (1984). Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-32(6)**:1109–1121.
- Ephraim, Y. und D. Malah (1985). Speech Enhancement Using a Minimum Mean-Square Error Log-Spectral Amplitude Estimator. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-33(2)**:443–445.
- Franklin, C. A., J. W. Thelin, A. K. Nabelek und S. B. Burchfield (2006). The Effect of Speech Presentation Level on Acceptance of Background Noise in Listeners with Normal Hearing. *Journal of the American Academy of Audiology*, **17**:141–146.
- Fredelake, S. (2006). *Untersuchung der perzeptiven Relevanz der Modulationstransferfunktion von Hörgerätealgorithmen*, (Master's thesis). Universität Oldenburg.
- Freyaldenhoven, M., P. Plyler, J. Thelin, M. Hedrick und S. Huck (2006a). Effect of Speech Presentation Level on Acceptance of Noise in Hearing Aid Users. In *International Hearing Aid Research Conference (ICHON 2006)*. Lake Tahoe, USA.

- Freyaldenhoven, M., P. Plyler, J. Thelin, M. Hedrick und S. Huck (**2006b**). Effect of Speech Presentation Level on Acceptance of Noise in Listeners with Normal and Impaired Hearing. In *International Hearing Aid Research Conference (ICHON 2006)*. Lake Tahoe, USA.
- Freyaldenhoven, M. C., D. Fisher Smiley, R. A. Muenchen und T. N. Konrad (**2006c**). Acceptable Noise Level: Reliability Measures and Comparison to Preferences for Background Sounds. *Journal of the American Academy of Audiology*, **17**:640–648.
- Freyaldenhoven, M. C., A. K. Nabelek, S. B. Burchfield und J. W. Thelin (**2005**). Acceptable Noise Level as a Measure of Directional Hearing Aid Benefit. *Journal of the American Academy of Audiology*, **16**:228–236.
- Freyaldenhoven, M. C., P. N. Plyler, J. W. Thelin, A. K. Nabelek und S. B. Burchfield (**2006d**). Acceptance of noise with monaural and binaural amplification. *Journal of the American Academy of Audiology*, **17**(9):659–666.
- Freyaldenhoven, M. C., P. N. Pyler, J. W. Thelin, A. K. Nabelek und S. B. Burchfield (**2006e**). The Effects of Ventig and Low-Frequency Gain Compression on Performance in Noise with Directional Hearing Instruments. *Journal of the American Academy of Audiology*, **17**:168–178.
- Gordon-Hickey, S. und R. E. Moore (**2007**). Influence of Music and Music Perception on Acceptable Noise Levels in Listeners with Normal Hearing. *Journal of the American Academy of Audiology*, **18**(5):417–427.
- Hagerman, B. und A. Olofsson (**2004**). A Method to Measure the Effect of Noise Reduction Algorithms Using Simultaneous Speech and Noise. *Acta Acustica united with Acustica*, **90**:356–361.
- Harkrider, A. W. und S. B. Smith (**2005**). Acceptable Noise Level, Phoneme Recognition in Noise, and Measures of Audiotry Efferent Acitivity. *Journal of the American Academy of Audiology*, **16**(8):530–545.
- Harkrider, A. W. und J. W. Tampas (**2006**). Differences in responses from the cochleae and central nervous systems of females with low versus high acceptable noise levels. *Journal of the American Academy of Audiology*, **17**(9):667–676.
- Hawkins, J. E. und S. S. Stevens (**1950**). The masking of pure tones and of speech by white noise. *Journal of the Acoustical Society of America*, **22**:6–13.
- Holube, I., S. Fredelake, J. Bitzer und M. Vlaming (**2007**). Erstellung eines Testsignals mit Sprachcharakteristik. DAGA. Stuttgart.
- Holube, I., M. Hansen, R. Schulz-Amling und S. Fredelake (**2005**). The Use of the Modulation Transfer Function to Describe the Performance of Hearing Aids. In *Hearing Aid Fitting*, 21st Danavox Symposium.
- Houtgast, T. und H. J. M. Steeneken (**1985**). A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *Journal of the Acoustical Society of America*, **77**(3):1069–1076.

- Huber, R. (2003). *Objective assessment of audio quality using an auditory processing model*, (Ph.D. thesis). Fachbereich Physik, Universität Oldenburg.
- Kalikov, D., K. Stevens und L. Elliot (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*, **61**:1337–1351.
- Kammeyer, K. D. und K. Kroschel (2002). *Digitale Signalverarbeitung - Filterung und Signalverarbeitung* (5. Aufl.). Stuttgart: B. G. Teubner Verlag.
- Kollmeier, B. und M. Wesselkamp (1997). Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment. *Journal of the Acoustical Society of America*, **102**:2412–2421.
- Kroschel, K. (2004). *Statistische Informationstechnik* (4. Aufl.). Berlin - Heidelberg - New York: Springer-Verlag.
- Kuk, F. K (2002). *Strategies for Selecting and Verifying Hearing Aid Fittings - Paired Comparisons as a Fine-Tuning Tool in Hearing Aid Fittings* (2 Aufl.). New York, Stuttgart: Thieme Verlag.
- Larsby, B. und S. Arlinger (1994). Speech recognition and just-follow-conversation tasks for normal-hearing and hearing-impaired listeners with different maskers. *Audiology*, **33**(3):165–176.
- Latzel, M., J. Kießling, S. Margof-Hackl, M. Dörscher, T. Wolter und R. Blum (2003). *Das JFC-Verfahren - Alternative oder Ergänzung zur Sprachaudiometrie*. Technical Report, Oldenburg: Haus des Hörens.
- Lozán, J. L. und H. Kausch (1998). *Angewandte Statistik für Naturwissenschaftler* (2., überarbeitete und ergänzte Auflage Aufl.). Berlin: Parey Verlag.
- Lurquin, P., C. Delacressonniere und A. May (2001). Examination of a Multi-Band Noise System. *The Hearing Review*, **8**(1):48–54,60.
- Magnusson, L., M. Karlsson und A. Leijon (2001). Predicted and measured speech recognition performance in noise with linear amplification. *Ear & Hearing*, **22**(1):46–57.
- Martin, R. (1993). An Efficient Algorithm to Estimate the Instantaneous SNR of Speech Signals. In *Proc. EUROSPEECH*, S. 1093–1096.
- Martin, R. (1994). Spectral Subtraction Based on Minimum Statistics. In *Proc. Euro. Signal Processing Conf. (EUSIPCO)*, S. 1182–1185.
- Martin, R. (2001). Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics. *IEEE Transactions on Speech and Audio Processing*, **9**(5):504–512.
- Martin, R. (2003). Statistical Methods for the Enhancement of Noisy Speech. In *International Workshop on Acoustic Echo and Noise Control (IWAENC)*. Kyoto, Japan.

- Marzinzik, M. (2000). *Noise Reduction Schemes for Digital Hearing Aids and their Use for the Hearing Impaired*, (Ph.D. thesis). Fachbereich Physik, Universität Oldenburg.
- Mueller, H. G., J. Weber und B. W. Y. Hornsby (2006). The Effects of Digital Noise Reduction on the Acceptance of Background Noise. *Trends in Amplification*, **10**(2):83–93.
- Nabelek, A. K. (2005). Acceptance of background noise may be key to successful fittings. *The Hearing Journal*, **58**(4):10–15.
- Nabelek, A. K., M. C. Freyaldenhoven, J. W. Tampas und S. B. Burchfield (2006). Acceptable noise level as a predictor of hearing aid use. *Journal of the American Academy of Audiology*, **17**(9):626–639.
- Nabelek, A. K., J. W. Tampas und S. B. Burchfield (2004). Comparison of speech in backgroundnoise with acceptance of background noise in aided and unaided conditions. *Journal of Speech and Hearing Research*, **47**(5):1001–1011.
- Nabelek, A. K., F. M. Tucker und T. R. Letowski (1991). Toleration of background noises: relationship with patterns of hearing aid use by elderly persons. *Journal of Speech and Hearing Research*, **34**:679–685.
- Neumann, J. und N. S. Jensen (2001). *Just Follow Conversation - Ericsholm Version (Report)*. Technical Report: Oticon A/S research center.
- Newman, C., B. Weinstein, G. Jacobson und G. Hug (1990). The Hearing Handicap Inventory for Adults: Psychometric adequacy and audiometric correlates. *Ear & Hearing*, **11**:430–433.
- Nilsson, M., S. D. Soli und J. A. Sullivan (1994). Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America*, **95**(2):1085–1099.
- Peissig, J. (1992). *Binaurale Hörgerätestrategien in komplexen Störschallsituationen*, (Ph.D. thesis). mathematisch naturwissenschaftlicher Fachbereich, Universität Göttingen.
- Peissig, J. und B. Kollmeier (1997). Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners. *Journal of the Acoustical Society of America*, **101**(3):1660–1670.
- Plomp, R. und A. M. Mimpen (1979). Speech-reception threshold for sentences as a function of age and noise level. *Journal of the Acoustical Society of America*, **66**(5):1333–1342.
- Punch, J., A. Joseph und B. Rakerd (2004). Most Comfortable and Uncomfortable Loudness Levels: Six Decades of Research. *American Journal of Audiology*, **13**:144–157.

- Ricketts, T. A. und B. W. Y. Hornsby (2005). Sound Quality Measures for Speech in Noise through a Commercial Hearing Aid Implementing "Digital Noise Reduction". *Journal of the American Academy of Audiology*, **16**(5):270–277.
- Rogers, D. S., A. W. Harkrider, S. Burchfield und A. Nabelek (2003). The influence of listener's gender on the acceptance of background noise. *Journal of the American Academy of Audiology*, **14**:372–382.
- Sachs, L. (2000). *Angewandte Statistik* (10., überarbeitete und aktualisierte Auflage Aufl.). Berlin: Springer Verlag.
- Saunders, G. H., T. H. Chisolm und H. B. Abrams (2005). Measuring hearing aid outcomes - Not as easy as it seems. *Journal of Rehabilitation Research & Development*, **42**(4):157–168.
- Saunders, G. H. und K. M. Cienkowski (2002). A Test to Measure Subjective and Objective Speech Intelligibility. *Journal of the American Academy of Audiology*, **13**:39–49.
- Saunders, G. H., A. Forsline und S. A. Fausti (2004). The Performance-Perceptual Test and Its Relationship to Unaided Reported Handicap. *Ear & Hearing*, **25**(2):117–126.
- Saunders, G. H., A. Forsline und S. A. Fausti (2006). The Performance-Perceptual Test and Its Relationship to Aided Reported Handicap. *Ear & Hearing*, **27**(3):229–242.
- Schick, A. (1979). *Schallwirkung aus psychologischer Sicht*. Klett Verlag.
- Sixtl, F. (1967). *Meßmethoden der Psychologie - Theoretische Grundlagen und Probleme*. Weinheim: Verlag Julius Beltz.
- Smiley, D. F., M. C. Freyaldenhoven, R. A. Muenchen und T. N. Konrad (2006). Acceptable noise level: reliability measures and comparison to preference of background noise. *Journal of the American Academy of Audiology*, **17**(9):640–648.
- Tchorz, J. und B. Kollmeier (2003). SNR Estimation Based on Amplitude Modulation Analysis With Applications to Noise Suppression. *IEEE Transactions on Speech and Audio Processing*, **11**(3):184–192.
- Vary, P., U. Heute und W. Hess (1998). *Digitale Signalverarbeitung*. Teubner Verlag.
- Ventry, I. und B. Weinstein (1982). The Hearing Handicap Inventory for Elderly: A new tool. *Ear & Hearing*, **3**:40–46.
- von Hapsburg, D. und J. Bangh (2006). Acceptance of Background Noise Levels in Bilingual (Korean-English) Listener. *Journal of the American Academy of Audiology*, **17**(9):649–658.
- Wagener, K. (2003). *Factors Influencing Sentence Intelligibility in Noise*, (Ph.D. thesis). Fachbereich Physik, Universität Oldenburg.

- Wagener, K., T. Brand und B. Kollmeier (**1999a**). Entwicklung und Evaluation eines Satztests für die deutsche Sprache Teil II: Optimierung des Oldenburger Satztests. *Zeitschrift für Audiologie*, **38**(2):44–56.
- Wagener, K., T. Brand und B. Kollmeier (**1999b**). Entwicklung und Evaluation eines Satztests für die deutsche Sprache Teil III: Evaluation des Oldenburger Satztests. *Zeitschrift für Audiologie*, **38**(3):86–95.
- Wagener, K., V. Kühnel und B. Kollmeier (**1999c**). Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests. *Zeitschrift für Audiologie*, **38**(1):4–15.
- Walden, B. E., R. K. Surr, M. T. Cord, B. Edwards und L. Olson (**2000**). Comparison of Benefits Provided by Different Hearing Aids Technologies. *Journal of the American Academy of Audiology*, **11**(10):540–560.
- Wittkop, T. (**1994**). *Vergleich binauraler digitaler Hörgerätestrategien zur Störgeräuschreduktion*. (Diploma thesis). Drittes Physikalisches Institut, Universität Göttingen.
- Wittkop, T. (**2001**). *Two-channel noise reduction algorithms motivated by models of binaural interaction*, (Ph.D. thesis). Fachbereich Physik, Universität Oldenburg.

Danksagung

Danken möchte ich meinen Prüfern Prof. Dr. Inga Holube, die mir das Thema dieser Masterarbeit überlassen hat, und Dr. Thomas Brand. Beide haben mir bei offenen Fragen geholfen und mich in jeder Phase der Arbeit sehr sachkundig und richtungsweisend begleitet und unterstützt.

Bedanken möchte ich mich auch bei Mitgliedern des Instituts Hörtechnik und Audiologie, für ihre Hilfe bei anstehenden Problemen.

Mein ganz besonderer Dank gilt auch den 84 Versuchspersonen, die mir in mehr als 160 Terminen für die Tests zur Verfügung gestanden haben.

Auch meiner Mutter, die unermüdlich Korrektur gelesen hat, immer wieder die richtigen Worte fand und an den richtigen Stellen die Ohren auf Durchzug gestellt hat, gilt mein herzliches Dankeschön.