

# Untersuchungen zum „Informational Masking“ bei unilateral und bimodal versorgten erwachsenen Cochlea-Implantat-Trägern

Inaugural-Dissertation  
zur  
Erlangung der Doktorwürde  
der  
Philosophischen Fakultät  
der  
Rheinischen Friedrich-Wilhelms-Universität  
zu Bonn

vorgelegt von  
**Verena Pyschny**  
aus  
Dinslaken

Bonn 2014

Gedruckt mit der Genehmigung der Philosophischen Fakultät der Rheinischen  
Friedrich-Wilhelms-Universität Bonn

**Zusammensetzung der Prüfungskommission:**

Professor Dr. Caja Thimm  
(Vorsitzende)

Professor Dr. Wolfgang Hess  
(Betreuer und Gutachter)

Professor Dr. Claudia Wich-Reif  
(Gutachterin)

apl. Professor Dr. Ulrich Schade  
(weiteres prüfungsberechtigtes Mitglied)

Tag der mündlichen Prüfung: 18.02.2014

## Danksagung

Ein ganz besonders herzlicher Dank geht an meinen Doktorvater, Herrn Prof. Dr. Wolfgang Hess, der mich während der Zeit meiner Promotion motivierte und durch Anregungen immer wieder unterstützt hat.

Während meiner Zeit am Jean-Uhrmacher-Institut (JUI) für klinische HNO-Forschung konnte ich die Studien im Rahmen des Projektes „Informational Masking and Speaker Identification in Cochlear-Implant-Recipients“ durchführen und daraus resultierend die folgende Arbeit erstellen. Deswegen geht mein Dank an Herrn Prof. Dr. Hartmut Meister unter dessen Leitung ich in seinem Team mit der Dissertation beginnen konnte. Das Projekt wurde mit Hilfe der finanziellen Unterstützung der Firma Cochlear Europe Ltd. umgesetzt.

Mein Dank gilt auch den Kollegen aus dem JUI und den Kollegen des Cochlear Implant Centrums Köln. Viele unterstützten mich auf die eine oder andere Weise. Besonders hervorheben möchte die Kollegen und Freunde Linda Müller und Markus Landwehr.

Vielen Dank an meine Probanden für die Zeit, Geduld und das Interesse an meiner Arbeit, sowie an die Korrekturleser Prof. Dr. Prisca Stenneken, Linda Müller und meine Mama.

Last but not least möchte ich mich an meine Eltern und meine Freunde wenden, die in der langen Schaffensphase hinter mir standen und mich immer motiviert haben und an meinen Schatz Til, der mir durch seine Ruhe und Gelassenheit manch stressigen Moment erleichterte.

# Inhaltsverzeichnis

Inhaltsverzeichnis	i
Abbildungsverzeichnis	v
Tabellenverzeichnis	vii
<b>1 Einleitung</b>	<b>1</b>
1.1 Motivation.....	2
1.2 Überblick.....	3

## ERSTER TEIL: THEORETISCHE GRUNDLAGEN

<b>2 Das auditorische System</b>	<b>6</b>
2.1 Anatomie des Gehörs.....	7
2.2 Physiologie des Gehörs .....	9
2.2.1 Reizübertragung .....	11
2.2.2 Sprach- und Sprechererkennung.....	12
2.2.3 Binaurales Hören.....	15
2.3 Pathophysiologie des Hörens.....	18
<b>3 Das Cochlea-Implantat</b>	<b>22</b>
3.1 Indikation.....	22
3.2 Aufbau und Funktionsweise eines Cochlea-Implantates.....	23
3.3 Kodierung akustischer Signale mit dem Cochlea-Implantat.....	26
3.4 Versorgungsarten.....	29
3.4.1 Bimodale und bilaterale Versorgung .....	30
3.4.2 Exkurs Hörgeräte.....	31
3.5 Möglichkeiten und Grenzen beim Hören mit dem Cochlea-Implantat .....	33
3.5.1 Sprechererkennung mit CI .....	34

3.5.2	Sprachverstehen im Störgeräusch mit CI.....	35
3.5.3	Prosodie- und Musikwahrnehmung mit CI .....	36
3.5.4	Beidohriges Hören mit Hörhilfen .....	38
<b>4</b>	<b>Sprachverstehen im Störgeräusch</b>	<b>40</b>
4.1	Periphere und zentrale Maskierung.....	41
4.1.1	Informational Masking - Konzept.....	43
4.1.2	„Stimulus uncertainty“ und „Target-masker similarity“ .....	44
4.1.3	Informational Masking mit Sprachsignalen.....	45
4.1.4	Exemplarische Untersuchungen mit Normalhörenden .....	47
4.1.5	Exemplarische Untersuchungen mit Cochlea-Implantat Trägern.....	50
<b>5</b>	<b>Zusammenfassung</b>	<b>53</b>
 <b>ZWEITER TEIL: EMPIRISCHER TEIL</b>		
<b>6</b>	<b>Zielsetzung und Hypothesen</b>	<b>56</b>
6.1	Vortest: Sprecherunterscheidung und Sprechererkennung .....	56
6.2	Experiment 1: Pegel- und Grundfrequenzunterschiede .....	57
6.3	Experiment 2: Grundfrequenz- und Formantfrequenzunterschiede .....	57
6.4	Experiment 3: Effekte räumlichen Hörens .....	59
<b>7</b>	<b>Vortest - Sprecherunterscheidung und Sprechererkennung</b>	<b>61</b>
7.1	Zielsetzung.....	62
7.2	Methode und Durchführung.....	62
7.3	Ergebnisse .....	64
7.3.1	Experiment 1: Sprecherunterscheidung .....	64
7.3.2	Experiment 2: Sprechererkennung .....	65
7.4	Diskussion.....	68
7.5	Fazit .....	69
<b>8</b>	<b>Methode und Durchführung</b>	<b>70</b>
8.1	Oldenburger Satztest .....	70
8.2	Oldenburger Satztest bei IM.....	71
8.3	Stimuli der Maskierungsexperimente.....	73
8.3.1	Experiment 1: Pegel- und Grundfrequenzunterschiede.....	73
8.3.2	Experiment 2: Grund- und Formantfrequenzunterschiede.....	74

8.3.3	Experiment 3: Effekte räumlichen Hörens .....	75
8.4	Probandenkollektiv .....	76
8.4.1	Experiment 1 .....	78
8.4.2	Experiment 2 .....	78
8.4.3	Experiment 3 .....	78
8.5	Durchführung .....	78
8.6	Hör- und Messkonditionen.....	79
8.7	Ethik.....	80
<b>9</b>	<b>Analyse und Statistik</b>	<b>81</b>
9.1	Fehleranalyse.....	81
9.2	Auswertung .....	82
9.2.1	Experiment 1 .....	82
9.2.2	Experiment 2 .....	82
9.2.3	Experiment 3 .....	83
<b>10</b>	<b>Ergebnisse zu Experiment 1: Pegel- und Grundfrequenzunterschiede</b>	<b>85</b>
10.1	Zielverständlichkeit.....	85
10.2	Fehlerverteilung .....	87
10.2.1	Anteil der A-Typ-Fehler.....	87
10.2.2	Anteil der B-Typ-Fehler.....	88
10.3	Diskussion.....	89
10.4	Zusammenfassung der Ergebnisse .....	91
<b>11</b>	<b>Ergebnisse zu Experiment 2: Grund- und Formantfrequenzunterschiede</b>	<b>92</b>
11.1	Zielverständlichkeit.....	92
11.2	Fehlerverteilung .....	94
11.2.1	Anteil der A-Typ-Fehler.....	95
11.2.2	Anteil der B-Typ-Fehler.....	96
11.3	Diskussion.....	98
11.4	Zusammenfassung der Ergebnisse .....	100
<b>12</b>	<b>Ergebnisse zu Experiment 3: Effekte räumlichen Hörens</b>	<b>102</b>
12.1	Zielverständlichkeit.....	102
12.2	Fehlerverteilung .....	106
12.3	Diskussion.....	108
12.4	Zusammenfassung der Ergebnisse .....	111

<b>13</b>	<b>Zusammenfassung und allgemeine Diskussion</b>	<b>112</b>
<b>14</b>	<b>Ausblick</b>	<b>117</b>
	<b>Literaturverzeichnis</b>	<b>120</b>
<b>A</b>	<b>Verzeichnisse</b>	<b>131</b>
A.1	Abkürzungen.....	131
A.2	Glossar.....	133
<b>B</b>	<b>Methodenbezogene Information</b>	<b>136</b>
B.1	Stimuligenerierung für den Vortest .....	136
B.2	Praat-Skript nach Darwin et al., 2003 .....	139
B.3	Basisliste des Oldenburger Satztests nach Wagener et al. (1999a) .....	141
B.4	Modifizierte Listen des Oldenburger Satztests für IM.....	142
B.5	Demografie der Probanden .....	145
B.6	Audiogramme der bimodal versorgten Probanden .....	148
<b>C</b>	<b>Statistische Auswertung</b>	<b>149</b>
C.1	Experiment 1 .....	149
C.2	Experiment 2.....	151
C.3	Experiment 3.....	153

## Abbildungsverzeichnis

2.1	Schematischer Aufbau des Ohres (Quelle: www.optik-lau-akustik.de).....	6
2.2	Seitlicher Längsschnitt der gewundenen Cochlea (Quelle: Netaudio, 2003)...	8
2.2.1	Querschnitt der Cochlea (Quelle: Netaudio, 2003) .....	8
2.2.2	Querschnitt des cortischen Organs (Quelle: Netaudio, 2003) .....	8
2.3	Einhüllende und Feinstruktur eines amplituden- und frequenz- modulierten Sinussignals (Quelle: Mitterbacher, 2004).....	12
2.4	Physikalische Unterschiede des Schallsignals bei Erreichen beider Ohren bei Schalleinfall von rechts (Quelle: Kießling et al., 2008).....	16
3.1	Interne (Quelle: Cochlear) und externe (Quelle: Med-EI) Komponenten eines Cochlea-Implantat-Systems - modifizierte Darstellung.....	24
3.2	Schematische Darstellung der Komponenten eines Cochlea-Implantates (Quelle: Med-EI) und Elektrodenlage in der Cochlea (Quelle: Cochlear).....	24
3.3	Darstellung der unterschiedlichen Stimulationsmodi (Quelle: McDermott, 2004) .....	25
3.4	Biphasisch ladungsausgeglichener Puls (Quelle: Hessel, 2004) .....	26
3.5	Blockdiagramm der CIS-Kodierungsstrategie (Quelle: Mitterbacher, 2004) ..	28
3.6	Schematische Darstellung des Aufbaus eines digitalen Hörgerätes (Quelle: Böhme und Welzl-Müller, 2005) .....	32
7.1	Ergebnisse für die Unterscheidung des Sprechergeschlechts .....	65
7.2	Ergebnisse für die Identifikation des Sprechergeschlechts	
7.2.1	Änderungen in F0.....	66
7.2.2	Änderungen in VTL .....	67
7.2.3	Änderungen in F0 und VTL .....	68
8.1	Darstellung der Maskiererarten im Zeitbereich.....	76
9.1	Hörkonditionen für die Berechnung der Effekte räumlichen Hörens .....	83



## **Abbildungsverzeichnis**

---

10.1	Zielverständlichkeit gemittelt über die 6 CI-Probanden.....	86
10.2	Anteil der A-Typ-Fehler gemittelt über die 6 CI-Probanden .....	87
10.3	Anteil der B-Typ-Fehler gemittelt über die 6 CI-Probanden .....	88
11.1	Durchschnittliches Sprachverstehen der 12 CI-Probanden .....	93
11.2	Durchschnittlicher Anteil der A-Typ-Fehler der 12 CI-Probanden.....	95
11.3	Durchschnittlicher Anteil der B-Typ-Fehler der 12 CI-Probanden.....	97
12.1	Differenz der Zielverständlichkeit (für CI) in Prozent .....	103
12.2	Differenz der Zielverständlichkeit (für HG) in Prozent .....	104
12.3	Differenz der A- und B-Typ-Fehler (für CI) in Prozent .....	106
12.4	Differenz der A- und B-Typ-Fehler (für HG) in Prozent.....	107
B.1	Stimmmodifizierung durch Änderungen von F0 und VTL (s. Tab. B.1).....	136
B.2	Stimmmodifizierung durch Änderungen von F0 und VTL (s. Tab. B.2).....	138
B.6.	Audiogramme der bimodal versorgten Probanden	
B.6.1	Experiment 2: Audiogramme der nicht-implantierten Seite .....	148
B.6.2	Experiment 3: Audiogramme der nicht-implantierten Seite.....	148

## Tabellenverzeichnis

7.1	Änderungen in F0 und in VTL bzw. $F_n$ (F1, F2, F3) in Hz für insgesamt 9 Stimuli für Experiment 2: Sprechererkennung.....	63
8.1	Werte für die Grundfrequenz und die Vokaltraktlänge mit einhergehenden Formantfrequenzen der Maskiererstimuli .....	74
B.1	Originalkategorien für die Änderungen von F0 und VTL (nach Darwin et al., 2003).....	136
B.2	Interpolierte Werte für das Sprecherunterscheidungsexperiment.....	137
B.3	Wortschatz und Konstruktion der Basisliste des Oldenburger Satztests.....	141
B.5	Demografie der Probanden	
B.5.1	Allgemeine Informationen zu den 6 unilateral versorgten CI-Probanden (Vortest und Experiment 1) .....	145
B.5.2	Allgemeine Informationen zu den 12 bimodal versorgten CI-Probanden (Experiment 2) .....	146
B.5.3	Allgemeine Informationen zu den 10 bimodal versorgten CI-Probanden (Experiment 3) .....	147
C.1	Statistische Auswertung: Experiment 1	
C.1.1	Einfluss der Pegel- und Grundfrequenzunterschiede auf die Zielverständlichkeit.....	149
C.1.2	Einfluss der Pegel- und Grundfrequenzunterschiede auf den Anteil der A-Typ-Fehler .....	150
C.1.3	Einfluss der Pegel- und Grundfrequenzunterschiede auf den Anteil der B-Typ-Fehler.....	150

C.2	Statistische Auswertung: Experiment 2	
C.2.1	Einfluss der Hörhilfen auf die Zielverständlichkeit und die Fehlertypen .....	151
C.2.2	Einfluss der Ziel-Maskierer-Konditionen auf die Zielverständlichkeit .....	152
C.2.3	Einfluss der Ziel-Maskierer-Konditionen auf den Anteil der A-Typ-Fehler .....	142
C.3	Statistische Auswertung: Experiment 3	
C.3.1	Einfluss der Effekte räumlichen Hörens auf die Zielverständlichkeit .....	153
C.3.2	Einfluss der Maskierer, Hörhilfe und F0 auf die Zielverständlichkeit beim Kopfschatteneffekt .....	153
C.3.3	Einfluss der Maskierer, Hörhilfe und F0 auf die Zielverständlichkeit beim Squelcheffekt .....	154
C.3.4	Einfluss der Maskierer, Hörhilfe und F0 auf die Zielverständlichkeit beim Summationseffekt .....	155
C.3.5	Einfluss der Effekte räumlichen Hörens auf den Anteil der Fehler ...	155
C.3.6	Einfluss der Maskierer, Hörhilfe und F0 auf den Anteil der Fehler beim Kopfschatteneffekt .....	156
C.3.7	Einfluss der Maskierer, Hörhilfe und F0 auf den Anteil der Fehler beim Squelcheffekt .....	156
C.3.8	Einfluss der Maskierer, Hörhilfe und F0 auf den Anteil der Fehler beim Summationseffekt .....	157

# Kapitel 1

## Einleitung

Schon von weitem hörte man dieses Geräusch, das nur sehr große Städte hervorbringen können, einen Ton, der aus allen Tönen zugleich besteht, aus Stimmengewirr und Tiergeschrei, aus Glockengeläut und Geldgeklimper, aus Kinderlachen und Hammerschlägen, aus Besteckklappern und dem Zuknallen Tausender von Türen, [...] – ein grandioses Rauschen, vom Leben selbst erzeugt

(Walter Moers, Die 13 ½ Leben des Käpt'n Blaubär).

In wenigen alltäglichen Situationen befindet man sich in einer ruhigen Umgebung. Im Straßenverkehr z.B. sind es laufende Motoren, hupende Autos oder quietschende Reifen, die Geräusche erzeugen. In einem Restaurant sind es das Besteckklappern, angeregte Unterhaltungen am Tisch, das Stühlerücken, zusätzliche Hintergrundmusik u. ä., die einen hohen Geräuschpegel erzeugen. In eben solchen Situationen ist es schwierig und auch anstrengend, eine Unterhaltung zu führen und jedes Wort seines Gesprächspartners zu verstehen. Um seinem Gegenüber in einer solchen Geräuschkulisse trotzdem noch folgen zu können, versucht man sich auf ihn zu „konzentrieren“. Automatisch ablaufende Prozesse des auditorischen Systems sorgen dafür, dass unerwünschte Hintergrundgeräusche abgemildert werden und die Stimme unseres Gegenübers herausgefiltert werden kann.

Diese Fähigkeit ist abhängig von der Art des Störgeräusches. Bei Alltagsgeräuschen, wie z.B. Autoverkehr, entsteht eine Verdeckung des Nutzsymbols durch die spektrale Energie des Störsymbols, so dass das Nutzsymbols unhörbar werden kann. Dies wird mit dem Begriff „Energetic Masking, EM“ beschrieben. Im Gegensatz dazu stehen Kommunikationssituationen mit mehreren, gleichzeitig sprechenden Personen. Hierbei entsteht die Verdeckung des Nutzsymbols nicht allein durch die spektrale Energie des Störsymbols. Das Gesagte des Störsprechers enthält zusätzliche Information. Diese Information ist jedoch irrelevant für den Hörer und hindert ihn daran, das Nutzsymbols zu erkennen und es vom Störsymbol zu unterscheiden. Der Hörer wird durch diesen simultanen Sprecher abgelenkt. Diese Art der Maskierung wird als „Informational Masking, IM“ bezeichnet.

Für Personen mit Hörstörungen sind besonders Hörsituationen mit mehreren Sprechern (IM), die mit dem Oberbegriff „Cocktail Party Effekt“<sup>1</sup> umschrieben werden, extrem schwierig und anstrengend. Die automatisch ablaufenden Prozesse eines normalhörenden Ohres sind nicht mehr in der gleichen Weise aktiv. Technische Hörhilfen wie Hörgeräte (HG) oder eine elektrische Innenohrprothese (Cochlea-Implantat, CI) können in Kommunikationssituationen mit derartigen Störgeräuschen nur bedingt Abhilfe schaffen. Unterschiedliche Algorithmen in Hörhilfen erkennen eine bestimmte Art von Störgeräusch und können diese vom modulierten Sprachsignal unterscheiden. Die sog. stationären Störgeräusche werden abgedämpft, so dass das Sprachsignal klarer herausgefiltert wird und besser wahrgenommen werden kann. Ist Sprache jedoch das Störgeräusch, greifen diese Algorithmen nicht.

### **1.1 Motivation**

Um das Verstehen in alltäglichen Hörsituationen einschätzen zu können und somit einen eventuellen Hörverlust diagnostizieren und behandeln zu können, werden sprachaudiometrische Messungen durchgeführt. In der Sprachaudiometrie gibt es eine Vielzahl von Sprachtests, die sich über das Verstehen von einzelnen Logatomen, Wörtern bis hin zu Sätzen erstrecken. Diese Tests werden entweder in Ruhe durchgeführt oder mit einem zusätzlichen Störgeräusch dargeboten, um das Hören der Patienten in möglichst vielfältigen Alltagssituationen zu repräsentieren.

Bei diesen Tests werden unterschiedliche Rauscharten – stationär oder moduliert – als Störgeräusch verwendet. Im deutschsprachigen Raum gibt es bis dato noch keinen sprachaudiometrischen Test, bei dem explizit *Sprache* als Störgeräusch genutzt wird. Somit fehlt die irrelevante Information als Störfaktor, die bei gleichzeitig dargebotener Sprache hinzukommt. Im Wesentlichen wird nur der Effekt des EM und nicht des IM untersucht. Aus wissenschaftlichen Studien mit Normalhörenden (NH) ist jedoch bekannt, dass die Verwendung von Rausch- oder Sprachsignalen als Störgeräusch unterschiedliche Auswirkungen auf das Sprachverstehen hat. Die Auswirkungen der Maskierung auf Hörgeschädigte sind nicht mit den Auswirkungen auf NH vergleichbar. Wie sich Sprache als Störgeräusch auf das Verstehen bei Trägern eines CIs auswirkt, ist nur unzureichend untersucht.

Effekte, die durch simultane Sprecher hervorgerufen werden, lassen sich höchstwahrscheinlich nicht hinreichend durch moduliertes Rauschen abbilden und spiegeln somit nicht die Probleme wider, denen Patienten im Alltag begegnen. Es ist notwendig, die

---

<sup>1</sup> Bezeichnet die Fähigkeit die Aufmerksamkeit selektiv zu steuern. Beispielsweise kann in einer lauten Umgebung einem Gesprächspartner zugehört werden, während Gespräche anderer Personen unterdrückt, aber weiterhin unbewusst wahrgenommen werden. Die Aufmerksamkeit kann sofort umgelenkt werden, wenn in einem anderen Gespräch der eigene Name genannt wird.

## **Kapitel 1. Einleitung**

---

Parameter zu kennen, die den CI-Trägern Vor- oder Nachteile in solchen Kommunikationssituationen erbringen.

In der vorliegenden Arbeit soll untersucht werden, inwieweit es für CI-Träger möglich ist, einen von zwei konkurrierenden Sprechern zu verstehen, und welche akustischen Parameter zu einer besseren Trennung und somit einer Reduzierung von IM beitragen.

Aspekte wie Pegelunterschiede zwischen den Sprechern und die unterschiedliche stimmliche Charakteristik der Sprecher werden dabei untersucht. Weiterführend werden nicht nur unilateral versorgte CI-Probanden an der Studie teilnehmen, sondern auch Probanden, die mit einem CI auf dem einen Ohr und mit einem HG auf dem kontralateralen Ohr versorgt sind (bimodales Hören). Durch das HG werden tieffrequente Anteile akustisch auf das Restgehör der Patienten abgebildet, was die Übertragung der stimmlichen Charakteristik, insbesondere die der Grundfrequenz ( $F_0$ ), verbessert. Dies könnte zu einer besseren Sprachverständlichkeit bei konkurrierenden Sprechern führen. Somit kann festgestellt werden, inwieweit das HG einen zusätzlichen Gewinn zum Sprachverstehen mit dem CI alleine erbringt. In der bimodalen Konstellation kann sich ebenfalls zeigen, ob Prozesse des binauralen Hörens greifen und Kommunikationssituationen mit konkurrierenden Sprechern erleichtert werden.

Allgemein sollen die Untersuchungen dazu dienen, annähernd die Schwierigkeiten darzustellen, denen CI-Träger im Gespräch mit mehreren, gleichzeitig sprechenden Personen im Alltag begegnen. Da es sich um sehr schwierige Kommunikationsszenarien handelt, werden ausschließlich erwachsene Probanden mit sehr gutem Sprachverstehen in die Studie einbezogen.

### **1.2 Überblick**

Das vorliegende Kapitel beschreibt den Themenbereich und die damit einhergehende Motivation für diese Arbeit. Im Theorieteil führt das zweite Kapitel in die Grundlagen des Hörens und die Verarbeitung von akustischen Signalen ein. Ein grundlegendes Wissen über die Akustik von (Sprach-)Signalen ist unerlässlich und wird für das Verstehen dieser Arbeit vorausgesetzt. Der Fokus von Kapitel 2 liegt vor allem auf dem Innenohr und dessen Funktionen. Die Innenohrfunktionen des pathologisch veränderten Gehörs werden mit den wesentlichen Funktionen des normalhörenden Ohres verglichen. Damit werden die Grundlagen gelegt, um die in Kapitel 3 beschriebene Verarbeitung von Signalen, wie sie durch ein CI übertragen werden, nachvollziehen zu können. Kapitel 3 beschäftigt sich mit der Funktionsweise eines CIs. Aufgrund des bimodal versorgten Probandenkollektivs in dieser Studie erfolgt auch ein kurzer Überblick über den Aufbau und die Funktion von Hörgeräten.

## **Kapitel 1. Einleitung**

---

Es werden die Möglichkeiten und Probleme erläutert, denen sowohl unilateral als auch bimodal versorgte CI-Träger im Alltag begegnen.

Daran anschließend werden in Kapitel 4 die Besonderheiten vorgestellt, die sich bei dem Verstehen von Sprache mit gleichzeitig dargebotenem Störsignal ergeben. Hierbei werden die Unterschiede und Auswirkungen von EM im Gegensatz zum IM betrachtet und ein Überblick über bisherige Studien mit NH und CI-Trägern gegeben. Die Grundlagen werden in Kapitel 5 kurz zusammengefasst.

Anschließend folgt der empirische Teil dieser Arbeit. In Kapitel 6 werden Zielsetzungen und Hypothesen aufgestellt. In einem Vortest, der in Kapitel 7 beschrieben wird, wurde die Fähigkeit von CI-Trägern getestet, eine Sprecherstimme dem richtigen Geschlecht zuzuordnen und sequentiell dargebotene Sprecher unterscheiden zu können. Die Ergebnisse dieses Vortests legen die Basis für die folgenden Experimente. Erst wenn es möglich ist, Sprecher zu erkennen und voneinander zu unterscheiden, sind die Voraussetzungen für das Verstehen von simultan dargebotenen Sprechern gegeben.

Das achte Kapitel beschreibt die Methodik, mit der IM in der vorliegenden Arbeit untersucht wurde. Daran anschließend wird diese in drei unterschiedlichen Experimenten eingesetzt. Das erste Experiment widmete sich den Fähigkeiten von CI-Trägern, simultane Sprecher anhand von Pegel- und Grundfrequenzunterschieden zu trennen. Das Probandenkollektiv bestand aus unilateral versorgten CI-Trägern. Am zweiten Experiment nahmen bimodal versorgte CI-Träger teil, um den Zugewinn des kontralateralen HGs zum CI untersuchen zu können. Auch hier lag das Hauptaugenmerk auf der Fähigkeit, konkurrierende Sprecher voneinander zu trennen. Hierfür wurden die stimmlichen Charakteristiken der Sprecher wie  $F_0$  und die Formantfrequenzen ( $F_n$ ) näher untersucht. Letztendlich erfolgte im dritten Experiment eine Untersuchung zur räumlichen Trennung von Ziel- und Maskierersprecher. Es wurden zweierlei Maskierer generiert, um die Ergebnisse bei der Verwendung von Rausch- und Sprachmaskierern vergleichen und somit eine Aussage über die Effekte von EM und IM treffen zu können. Auch hier war das Probandenkollektiv bimodal mit Hörhilfen versorgt.

In Kapitel 9 werden die verwendeten Analyse- und Statistikmethoden vorgestellt, die anschließend in den Kapiteln 10, 11 und 12 auf die Ergebnisse der jeweiligen Experimente angewendet wurden. Die Ergebnisse werden jeweils mit einer kurzen Wiederholung der Methodik eingeleitet, um eine bessere Präsenz der Untersuchungen zu gewährleisten. Die Ergebnisse werden in Bezug auf die bisherige Literatur diskutiert. Daran anschließend folgt in Kapitel 13 eine Zusammenfassung mit einer allgemeinen Diskussion. Ein Ausblick auf weiterführende Forschungsfragestellungen wird in Kapitel 14 gegeben.

## **Kapitel 1. Einleitung**

---

Da diese Arbeit interdisziplinär die Gebiete der Linguistik, der Audiologie und der Medizin verbindet, werden in einem Glossar die grundlegenden Begrifflichkeiten beschrieben.



## Kapitel 2

### Das auditorische System

Das Gehör stellt die Grundlage für das Verstehen von Sprache dar und dient somit der menschlichen Kommunikation. Das menschliche Ohr ist theoretisch in der Lage, Frequenzen in einem Bereich von 20 Hertz (Hz) bis 20 Kilohertz (kHz) wahrzunehmen. Der Bereich zwischen Hörschwelle und Schmerzschwelle umfasst eine Dynamik von ca. 140 dB SPL (Schalldruckpegel: engl. sound pressure level (SPL) in Dezibel (dB)).<sup>2</sup>

Diese akustische Information wird vom menschlichen Hörorgan durch Außen-, Mittel- und Innenohr verarbeitet (s. Abbildung 2.1).

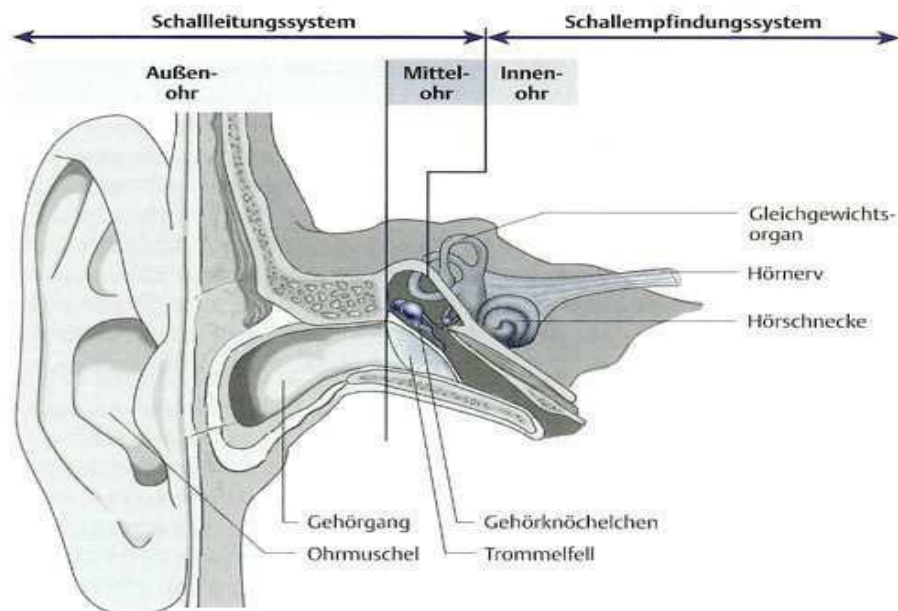


Abbildung 2.1: schematischer Aufbau des Ohres (Quelle: [www.optik-lau-akustik.de](http://www.optik-lau-akustik.de); 18.02.2013)

<sup>2</sup> Die nachfolgende Beschreibung der Anatomie und Physiologie des Hörsystems ist grundlegender Natur. Für detailliertere Information sei auf Hellbrück (2004) und Moore (2003) verwiesen. Soweit nicht anders gekennzeichnet, beziehen sich die Ausführungen auf diese beiden Bücher.

## **Kapitel 2. Das auditorische System**

---

An das Innenohr schließt der Hörnerv an, der als Teil der zentralen Hörbahn die Information an den auditorischen Kortex weiterleitet. Im auditorischen Kortex wird die Information weiterverarbeitet.

Das Hörsystem kann jedoch aufgrund seiner Komplexität in vielerlei Hinsicht in seiner Funktion gestört werden, so dass die Verarbeitung akustischer Signale stark eingeschränkt ist. In vielen Fällen kann das hörgeschädigte Ohr die Funktionen nicht mehr von selbst wiederherstellen, so dass nur noch eine apparative Therapie z.B. mit Hilfe einer Hörgeräteversorgung oder einer Versorgung mit einer Innenohrprothese (Cochlea-Implantat, CI) Abhilfe schaffen könnte.

### **2.1 Anatomie des Gehörs**

Das Außenohr besteht aus Ohrmuschel und Gehörgang und ist durch das Trommelfell vom Mittelohr abgegrenzt. Die Ohrmuschel ist der äußerste Teil des Ohres und besteht größtenteils aus elastischem Knorpel, der von Haut überzogen ist.

Der daran anschließende Gehörgang hat eine Länge von 3,0 - 3,5 cm und einen Durchmesser von 6 - 8 mm. Der äußere Teil des Gehörgangs enthält Haare mit Talgdrüsen, der innere Teil des äußeren Gehörgangs enthält dagegen keine Haare (Lenarz und Boenninghaus, 2012).

An den Gehörgang grenzt das Trommelfell, welches das Außen- vom Mittelohr abgrenzt. Daran anschließend befinden sich die miteinander verbundenen Gehörknöchelchen Hammer, Amboss und Steigbügel, die in der Paukenhöhle liegen. Die Paukenhöhle des Mittelohres ist ein mit Luft gefüllter Hohlraum, der das Außen- vom Innenohr abgrenzt und über die Eustachische Röhre mit dem Nasenrachenraum verbunden ist. Dem Mittelohr zugehörig sind auch zwei Binnenohrmuskel. Einer setzt am Hammergriff an, der andere am Steigbügelkopf (Lenarz und Boenninghaus, 2012). Die beweglichen Gehörknöchelchen verbinden das Trommelfell mit dem ovalen Fenster des Felsenbeins.

An das ovale Fenster schließt das Innenohr an, welches mit Flüssigkeit (Lymphe) gefüllt ist. Das Innenohr ist in den härtesten Knochen des menschlichen Körpers eingelagert, dem Felsenbein (Hoth, 1999). Es besteht aus Hohlräumen, die miteinander verbunden sind (Labyrinth). Dazu gehören sowohl die Bogengänge und der Vorhof (Vestibulum), die das Gleichgewichtsorgan bilden, als auch die Hörschnecke (Cochlea). Abbildung 2.2 verdeutlicht den Aufbau der Cochlea. Sie ist ein schneckenförmiges Rohr mit 2,5 Windungen und einer Länge von ca. 32 mm. Sie windet sich um die knöcherne Schneckenachse (Modiolus), welche die Ganglienzellen des Hörnervs enthält (Böhme und Welzl-Müller, 2005).

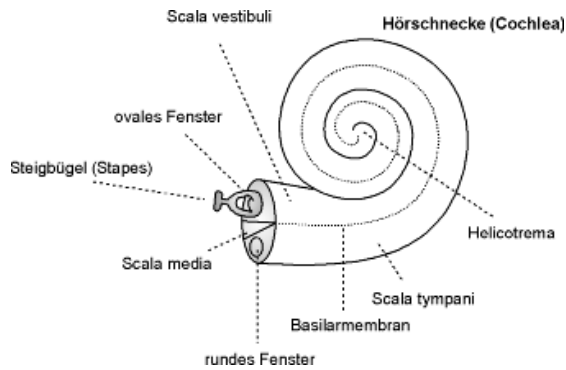
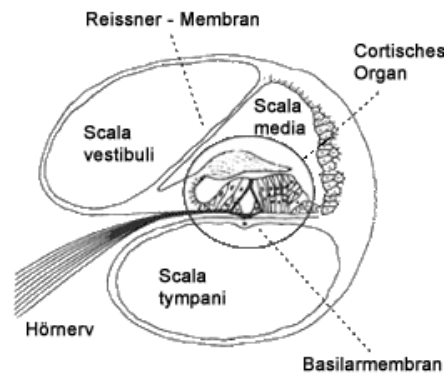
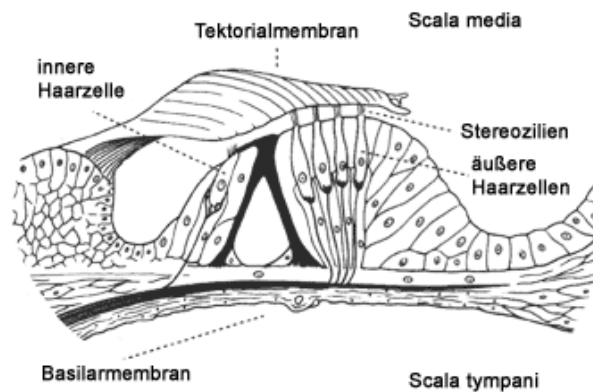


Abbildung: 2.2: Seitlicher Längsschnitt der gewundenen Cochlea



2.2.1: Querschnitt der Cochlea



2.2.2: Querschnitt des cortischen Organs (Quelle: Netaudio, 2003)

Die Schneckensbasis, an der sich das ovale und das runde Fenster befinden, wird als basales Ende bezeichnet. Die Schneckenspitze, das Helicotrema, wird als apikales Ende bezeichnet (s. Abbildung 2.2). Im Verlauf vom basalen zum apikalen Ende verringert sich der Durchmesser von 0,9 mm auf 0,3 mm. Die Cochlea ist in drei Hohlräume, die Scala vestibuli, die Scala tympani und die Scala media aufgeteilt. Die Scala vestibuli verläuft vom ovalen Fenster zum Helicotrema und ist dort mit der Scala tympani verbunden. Die Scala tympani erstreckt sich vom Helicotrema bis zum runden Fenster, welches diese Scala zum Mittelohr abgrenzt. Diese beiden Scalen sind mit Perilymphe gefüllt. Zwischen den beiden Scalen liegt die Scala media, die mit Endolymphe gefüllt ist. Sie ist durch die Reissnersche Membran von der Scala vestibuli und durch die Basilarmembran von der Scala tympani abgegrenzt (s. Abbildung 2.2.1). Die Basilarmembran wird vom basalen zum apikalen Ende hin immer breiter. Auf ihr sitzt das eigentliche Hörorgan, das Cortische Organ (Pompino-Marschall, 1995), in dem sich die Rezeptor- oder Haarzellen befinden (s. Abbildung 2.2.2). Es gibt eine Reihe innerer Haarzellen (ca. 3400) und drei Reihen äußerer Haarzellen (ca. 13400) (Böhme und Welzl-Müller, 2005), die in so genannte Stützzellen eingelagert sind. Auf den insgesamt ca. 17000 Haarzellen sitzen die Sinneshärchen (Stereozilien). Die Stereozilien der äußeren

Haarzellen berühren die Tektorialmembran, die das cortische Organ überdeckt. Die Stereozilien der inneren Haarzellen haben keinen Kontakt zur Tektorialmembran.

An den Haarzellen beginnen die Nervenfasern (30 000 Stk.) des Hörnervs. Die meisten Nervenfasern des Hörnervs sind afferent, d.h. Information wird von den Sinneszellen an das Gehirn weitergeleitet. Die afferenten Nervenfasern sind größtenteils mit den inneren Haarzellen verbunden (Hoth, 1999). Efferente (absteigende) Fasern, die Impulse vom Gehirn zur Peripherie leiten, ziehen hauptsächlich zu den äußeren Haarzellen und bedingen die Anpassung des peripheren Hörsystems an die entsprechende Hörumgebung (Lennarz und Boenninghaus, 2012).

Beide Arten von Nervenfasern ziehen von den Haarzellen zu den Spiralganglienzellen im Modiolus und verbinden sich dort zum Hörnerv. Von dort aus werden die auditorischen Reize über unterschiedliche Neuronen (Hörbahn) an den auditorischen Kortex weitergeleitet und verarbeitet. Die Hörbahn ist beim bewussten Hören für die Verarbeitung des Reizes bis hin zu seiner Wahrnehmung aktiviert (Lennarz und Boenninghaus, 2012).

### **2.2 Physiologie des Gehörs**

Die Aufgabe des Außenohres ist die Aufnahme und die Weiterleitung des Schalls aus der Umwelt zum Mittelohr (Pompino-Marschall, 1995). Die Schallweiterleitung ist richtungs- und frequenzabhängig und ist somit ein essentieller Faktor der Hörempfindung. Durch seine Resonanzeigenschaften werden im Gehörgang besonders Frequenzen im Bereich von 2 kHz bis 5 kHz verstärkt. Maximale Dämpfungen entstehen im Bereich von 8 kHz bis 10 kHz.

Außerdem besitzt das Außenohr eine Schutzfunktion. Fremdstoffe können nur schlecht eindringen, da der Gehörgang von außen nach innen immer schmaler wird. Zusätzlich kann der Ohrschmalz eventuell eingedrungene Fremdstoffe binden. Der Gehörgang und das Trommelfell trocknen aufgrund des Ohrschmalzes nicht aus, so dass das Trommelfell seine Elastizität und das Schwingungsverhalten beibehalten kann.

Das Mittelohr dient der Signalübertragung zwischen dem mit Luft gefüllten Außenohr und dem mit Lymphe gefüllten Innenohr. Die vom Außenohr ankommenden akustischen Schalldrücke werden über das Trommelfell und die Gehörknöchelchenkette an das ovale Fenster der Cochlea im Innenohr weitergeleitet. Zum Ausgleich des Drucks im Mittelohr mit der Umgebung dient die Eustachische Röhre, die mit dem Rachenraum verbunden ist.

Aufgrund der unterschiedlichen akustischen Eigenschaften von Luft (im Außen- und Mittelohr) und Flüssigkeit (im Innenohr) und des dadurch entstehenden unterschiedlichen Widerstandes, muss der Schall verstärkt werden. Der Druck am Trommelfell muss gegenüber dem Druck am ovalen Fenster erhöht werden. Diese Anpassungsfunktion übernimmt das Mittelohr. Das Trommelfell verfügt über eine größere Fläche als das ovale

## **Kapitel 2. Das auditorische System**

---

Fenster. Diese unterschiedlichen Flächenverhältnisse sorgen für eine Verstärkung. Zum anderen dient die Gehörknöchelchenkette als Hebelsystem, so dass eine größere Kraft auf das ovale Fenster einwirkt, als am Trommelfell ankommt. Ohne eine derartige Verstärkung würde der einkommende Schall zu 98% am Trommelfell reflektiert werden und nur 2% der Schallenergie würde tatsächlich ins Innenohr gelangen. Durch die Verstärkungseigenschaften des Mittelohres wird 60% der Energie an das Innenohr weitergegeben (Hoth, 1999).

Die Kontraktion der zwei Binnenohrmuskeln dient der Pegelanpassung und schützt das Innenohr vor zu hohen, andauernden Schallpegeln (Probst, 2008).

Im Innenohr wird die akustische Information in Nervenimpulse umgewandelt. Durch die Bewegung am ovalen Fenster kommt es zu Flüssigkeitsverschiebungen bzw. Druckänderungen in der Cochlea. Die Schwingungen können nur durch den gleichzeitigen Druckausgleich, der über das runde Fenster stattfindet, übertragen werden (Probst, 2008). Durch die Verschiebung kommt es zur Auslenkung der Basilarmembran in Form einer Wanderwelle, wie es 1960 von Békésy beschrieben wurde. Der Ort der Auslenkung der Wanderwelle auf der Basilarmembran ist frequenzabhängig. Die Beschaffenheit der Basilarmembran trägt dazu bei, dass Signale verschiedener Frequenz an unterschiedlichen Orten zu einer maximalen Auslenkung führen. Am basalen Ende ist die Basilarmembran steif, wohingegen sie am apikalen Ende flexibler ist. Hohe Frequenzen führen zu einer maximalen Auslenkung der Basilarmembran am basalen Ende (Basis) in der Nähe des ovalen Fensters. Bei tiefen Frequenzen erfährt die Basilarmembran die maximale Auslenkung im apikalen Bereich (Apex) in der Nähe des Helicotrema. Diese tonotope Zuordnung wird auch als Frequenz-Orts-Transformation bezeichnet.

Durch die Wanderwelle werden Basilar- und Tektorialmembran gegeneinander verschoben. Diese Verschiebung und die Druckveränderungen der umgebenden Flüssigkeit stimulieren die Stereozilien der äußeren Haarzellen. Die äußeren Haarzellen tragen aktiv zur Modulation der Wanderwelle bei. Sie wirken als Signalverstärker bei geringen Lautstärken und als Dämpfer bei hohen Lautstärken. Die äußeren Haarzellen präzisieren die Auslenkbewegungen der Basilarmembran und führen somit zu einer genaueren Frequenzanalyse.

Die durch die äußeren Haarzellen veränderte Bewegung der Basilarmembran führt zur Stimulierung der inneren Haarzellen. Die Auslenkung der Stereozilien der inneren Haarzellen führt zu Stoffwechselfvorgängen, durch die Neurotransmitter freigesetzt und Aktionspotentiale in den Hörnervenfasern ausgelöst werden (Wolf, 2004). Das eigentliche sensorische Signal wird produziert. Über die dadurch aktivierten Hörnervenfasern werden die Signale über die Hörbahn zum auditorischen Kortex weitergeleitet (Lenarz und Boenninghaus, 2012).

### **2.2.1 Reizübertragung**

Die tonotope Zuordnung tritt nicht nur in der Cochlea, sondern auch im Hörnerv und der weiter aufsteigenden Hörbahn auf. Ähnlich der Basilarmembran haben die stimulierten Nervenfasern eine charakteristische Frequenz. Die Schwelle für die Entladungen einer bestimmten Nervenfasern ist für diese charakteristische Frequenz am niedrigsten und steigt für umgebende Frequenzen an. Bei niedrigen bis mittleren Signalpegeln sind die Frequenzbereiche, auf die die Nervenfasern reagieren, sehr eng und somit die Frequenzselektivität präzise. Je höher jedoch der Signalpegel ist, desto breiter werden die Frequenzbereiche, auf die eine Reaktion der Nervenfasern erfolgt, und desto ungenauer ist die Frequenzselektivität.

Die Hörnervenfasern können weiterhin Information über das zeitliche Muster eines Signals übertragen und auswerten. Die Ausschüttung von Transmitterstoffen der inneren Haarzellen ist abhängig von der Phase der Auslenkung ihrer Stereozilien. Dies beeinflusst die neuronalen Impulse (Entladungen) der Nervenfasern (Wolf, 2004). Das Muster der Impulse folgt somit der Frequenz des ankommenden Signals. Dieser Effekt wird als Phasenkopplung bezeichnet. Das zeitliche Auflösungsvermögen der einzelnen Nervenfasern ist jedoch begrenzt, da die Nervenfasern nach einem abgegebenen Impuls eine Erholungsphase (Refraktärzeit) benötigen. Somit ist die Frequenzübertragung durch Phasenkopplung eher auf tiefere Frequenzen begrenzt. Sowohl die tonotope Zuordnung als auch die zeitliche Kodierung tragen zur Frequenzselektivität des Gehörs bei. Allerdings sind nicht beide Mechanismen gleichzeitig zu gleichen Anteilen aktiv. Während die zeitliche Kodierung eher im unteren Frequenzbereich (bis  $\sim 1$  kHz) greift, sind im mittleren Frequenzbereich beide Mechanismen aktiv. Ab  $\sim 5$  kHz werden die Informationen eher durch die tonotope Zuordnung übertragen.

Die Lautstärkeempfindung wird durch die Entladungsrate und durch die Anzahl der einzelnen aktiven Nervenfasern kodiert. Bei ansteigendem Signalpegel erhöht sich die Entladungsrate einer Nervenfasern. Ab einem gewissen Punkt kann sich die Entladungsrate nicht mehr erhöhen, so dass benachbarte Nervenfasern miterregt werden (Wolf, 2004). Allgemein ist die Übertragung von Signalen durch die Nervenfasern frequenz- und pegelabhängig. Die zeitliche Länge eines Signals wird über die Dauer der Aktivierung der Nervenzellen kodiert (Wolf, 2004).

Besonders für die Kodierung von Sprachsignalen ist die Zerlegung des Signals in zwei Bestandteile wichtig. Hilbert konnte 1912 die langsam variierende Einhüllende (Amplitudenmodulation) von der Feinstruktur, einer hochfrequenten Trägerschwingung mit gleichbleibender Amplitude (Frequenzmodulation), unterscheiden (s. Abbildung 2.3).

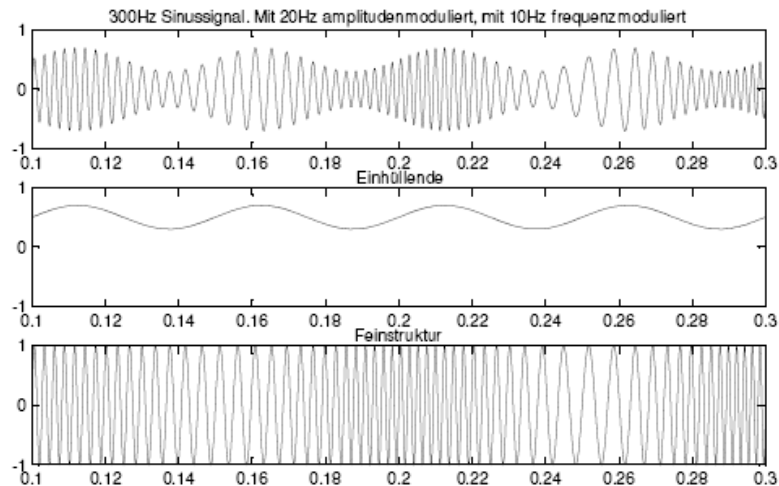


Abbildung 2.3: Einhüllende und Feinstruktur eines amplituden- und frequenzmodulierten Sinussignals  
(Quelle: Mitterbacher, 2004: 73)

Die Einhüllende trägt die Hauptinformation für nicht tonale sprachliche Äußerungen und ist ausschlaggebend dafür, dass Sprache verstanden wird. Die Feinstruktur hingegen bestimmt die Klangwahrnehmung (z.B. in der Musik) und erleichtert die Schalllokalisierung und das Sprachverstehen im Störgeräusch (Smith et al., 2002). Physiologisch gesehen geben die Neuronen aufgrund ihres zeitlichen Reaktionsmusters die Feinstruktur und aufgrund des Amplitudenmusters die Information der Einhüllenden eines Signals wieder (Med-EI, 2011).

### 2.2.2 Sprach- und Sprechererkennung

Der für die Sprachwahrnehmung wichtige Frequenzbereich liegt zwischen ca. 70 Hz und 8 kHz und umfasst einen Dynamikbereich von 40 dB SPL (zwischen 30 dB SPL und 70 dB SPL). Eine rein akustische Beschreibung des Sprachsignals reicht jedoch nicht aus, um dessen Wahrnehmung zu erklären. Der Wahrnehmung von Sprache unterliegen hochkomplexe Analysemechanismen, die weit über die periphere Verarbeitung und die Wahrnehmung der Tonhöhe, der Lautstärke und zeitlicher Muster des akustischen Signals hinausgehen. Diese Mechanismen werden durch die Segmenterkennung ergänzt, so dass das akustische Signal in Einheiten wie Laute, Silben und Wörter zusammengefasst werden kann (Wendler et al., 2005).

Alles, was wir hören, ist entweder Sprache oder keine Sprache – dazwischen scheint es nichts zu geben. Auch wenn wir uns noch so anstrengen, sind wir außerstande, eine gesprochene Äußerung als Abfolge von Summ- und Zischlauten zu hören. Was wir hören, sind Sprachlaute (Crystal 1998: 145).

Es gibt eine Vielzahl von Theorien, die sich mit der Verarbeitung von Sprache befassen. Die Spannweite reicht von einer rein auditiven Wahrnehmung bis hin zu speziellen

## Kapitel 2. Das auditorische System

---

Signalverarbeitungsmechanismen für Sprache. Bregman (1990) beschrieb mithilfe der auditiven Szenenanalyse, wie es dem auditorischen System gelingt, sinnvolle Einheiten zu erkennen und von anderen zu unterscheiden. Durch die sog. sequentielle Gruppierung werden aufeinanderfolgende Schallsignale wahrgenommen und verbunden. Die simultane Gruppierung beschreibt die Fähigkeit, gleichzeitig auftretende Schallsignale zu erkennen und voneinander zu trennen. Demnach bildet das auditorische System Gruppen oder Objekte, die zu einem auditiven Strom zusammengefasst werden.

Unterschiedliche Merkmale können als einem Objekt zugehörig identifiziert werden. In Bezug auf Sprache ist es vor allem die spezifische zeitliche und spektrale Struktur, durch die ein Objekt gebildet werden kann. Der gleichzeitige Beginn eines Ereignisses oder gleiche Amplitudenmodulation führen zur Wahrnehmung eines Objektes, das von einer bestimmten Signalquelle produziert wird. Auch die für den Sprachfluss typischen Fluktuationen, Unregelmäßigkeiten und Pausen werden als bedeutungstragend erkannt. Hörer sind in der Lage, Laute, Silben und Wörter auch über diese Unregelmäßigkeiten hinweg zu einer Phrase zu gruppieren. Hierbei spielen weitere Merkmale wie der Klang, die Grundfrequenz und der Ort der Schallquelle eine Rolle. Die Gruppierung unregelmäßiger Signalkomponenten im Zeitverlauf wird als „Streaming“ bezeichnet (Shinn-Cunningham und Best, 2008).

Bei der simultanen Gruppierung geht es darum, einzelne Komponenten aus einem komplexen Schallsignal zu extrahieren und zu einem Objekt der Wahrnehmung zu verbinden. Es werden Objekte gebildet, die zu einer Signalquelle zugehörig erscheinen und von anderen unterschieden werden können. In einer Situation mit mehreren Sprechern dienen hier vor allem sprecherspezifische Merkmale der Objektbildung. Im Anschluss sorgen Top-down Prozesse dafür, dass die Aufmerksamkeit selektiv auf eben diese Merkmale gelenkt wird. Es finden Verarbeitungsprozesse statt, die dem Abgleich mit sprachlichen (semantisch, syntaktisch, lexikalisch) Strukturen im mentalen Lexikon dienen. Diese Prozesse und das allgemeine Wissen des Hörers resultieren im Verstehen von Sprache.

In einer komplexen Hörsituation kann es sein, dass “a listener may hear properly formed words (due to shortterm object formation) and sentences (due to streaming), but still may have trouble selecting which stream to attend” (Shinn-Cunningham und Best, 2008: 287). In dieser Situation hilft dem Hörer Vorabinformation über den Inhalt des Gesagten und über den Sprecher.

Einerseits gelten die Grundfrequenz ( $F_0$ ) und die Formantfrequenzen ( $F_n$ ) als Träger linguistischer Information, durch die Sprachsignale gruppiert und verstanden werden können. Andererseits kann durch  $F_0$  und  $F_n$  auf die individuellen Eigenschaften eines Sprechers wie Geschlecht, Alter, Größe etc. geschlossen werden.

Die Höhe bzw. Tiefe von  $F_0$  wird durch die Länge der Stimmbänder bestimmt. Die Stimmbänder männlicher Sprecher sind im Mittel etwa 17 cm lang, wohingegen die



## Kapitel 2. Das auditorische System

Stimmbänder weiblicher Sprecher ungefähr 13 cm lang sind. Lange Stimmbänder schwingen langsam und bringen eine tiefe  $F_0$  hervor, wohingegen kurze Stimmbänder schnell schwingen und somit eine höhere  $F_0$  hervorbringen. Die typische  $F_0$  eines Mannes liegt bei ca. 100 Hz, die  $F_0$  einer Frauenstimme liegt ungefähr 1 Oktave höher bei 200 Hz (Peterson und Barney, 1952).

Vokale und andere stimmhafte Laute (Nasale) sind durch die unterschiedliche Lage von  $F_n$  charakterisiert. Jeder Vokal hat eine charakteristische Formantlage, die sich nur wenig mit  $F_0$  verändert. Das Verhältnis der  $F_n$  zueinander bleibt in etwa gleich, unabhängig vom Sprechenden. Für die Vokalerkennung und -differenzierung sind besonders die ersten beiden Formanten wichtig. Die  $F_n$  geben jedoch nicht nur Auskunft über den gesprochenen Vokal, sondern lassen Rückschlüsse auf den Vokaltrakt (primär Mund- und Rachenraum, sekundär Nasenraum) des Sprechenden zu. Als sprecherspezifisch gelten besonders die höheren  $F_n$ . Insgesamt gilt, dass  $F_n$  tiefer sind, je länger der Vokaltrakt ist. Der Vokaltrakt von Frauen ist im Mittel 16% kürzer als der von Männern, was in höheren  $F_n$  resultiert (Peterson und Barney, 1952).

Neben sprecherspezifischen Eigenschaften wie Sprechgeschwindigkeit, Rauigkeit, oder Behauchtheit der Stimme, können unterschiedliche Sprecher, und besonders Sprecher unterschiedlichen Geschlechts, anhand von  $F_0$  und  $F_n$  erkannt und unterschieden werden.

Weiterhin zeigt sich in Experimenten zum Sprachverstehen im Störgeräusch vor allem die Wichtigkeit der Wahrnehmung von  $F_0$ . Zum einen können durch  $F_0$  stimmlose von stimmhaften Lauten unterschieden werden, was dazu führt, dass weniger Wörter im mentalen Lexikon aktiviert werden (Meister et al., 2011b). Zum anderen dient  $F_0$  als Parameter, durch den Hörer in der Lage sind, ein Sprachsignal zu „streamen“, was für die Trennung konkurrierender Sprachsignale von großer Bedeutung ist (Assmann und Summerfield 1990).

$F_0$  ist weiterhin ein wichtiger Faktor bzgl. des sog. „Lückenhörens“ oder engl. „glimpsing“. Glimpsing bezieht sich auf die Fähigkeit Normalhörender bei konkurrierenden Signalen weniger stark maskierte Bereiche zu nutzen, um ein Zielsignal zu verstehen. Normalhörende können spektrale und temporale Bereiche ausnutzen, in denen das Sprachsignal energiereicher ist als das Störgeräusch (Cooke, 2005). Sprache kann somit in einem fluktuierenden Störgeräusch identifiziert und bis zu einem gewissen Grad verstanden werden. „The assumption is that listeners are able to first detect “useful” glimpses of speech, possibly occurring at different times and occupying different regions of the spectrum, and then somehow integrate those glimpses to hear out the target speech” (Li und Loizou, 2007: 1166).  $F_0$  ist ein Merkmal, das genutzt werden kann, um das Zielsignal in einer Menge von überlagerten Signalen zu erkennen. Sie zeigt somit an, wann in vorteilhafte zeitliche Lücken eines fluktuierenden Maskierers gehört werden kann (Brown und Bacon, 2009). Zusätzlich

zu  $F_0$  sind auch  $F_n$  und die Energie des Signals nutzbare Komponenten, um das Zielsignal zu „glimpsen“.

Außerdem ist  $F_0$  wichtig für die Wahrnehmung prosodischer Merkmale (Lehiste, 1970). Durch unterschiedliche Betonungen können sich unterschiedliche Bedeutungen von Wörtern ergeben, Frage- und Aussagesätze diskriminieren oder auch wichtige Aspekte im Satz hervorheben lassen. Auch die Erkennung von Wortgrenzen basiert auf der Wahrnehmung von  $F_0$ . Die „prosodische“ Funktion der  $F_0$  ist beim Sprachverstehen im Störgeräusch ebenfalls von großer Bedeutung.

Binns und Culling (2007) untersuchten die Funktion von Intonationskonturen in komplexen Hörsituationen. Sätze eines Zielsprechers wurden entweder mit regulärer  $F_0$ , mit verringerten Variationen in  $F_0$  oder mit invertierter  $F_0$ -Kontur dargeboten. Die Invertierung hatte zur Folge, dass sekundäre Elemente der Sprache wie Funktionswörter betont und Hauptinformationsträger wie Verben und Nomen unbetont waren. Der Zielsprecher war entweder durch einen weiteren Sprecher oder durch ein sprachsimulierendes Rauschen verdeckt. Bei Verwendung des sprachsimulierenden Rauschens veränderte sich das Sprachverstehen durch die abgeflachte  $F_0$ -Kontur nur wenig. Die Invertierung von  $F_0$  führte jedoch zu einer Verschlechterung. Bei Verwendung eines Sprechers als Störsignal sank das Sprachverstehen weiter, wobei auch hier die Invertierung zu einem deutlich schlechteren Sprachverstehen führte als die abgeflachte  $F_0$ -Kontur. Eine Manipulation der Maskiererstimme bzgl. der  $F_0$ -Kontur veränderte hingegen weder die Maskierungseigenschaften noch das Sprachverstehen.

### **2.2.3 Binaurales Hören**

Mit der Bezeichnung „binaurales Hören“ wird die Fähigkeit beschrieben, die an beiden Ohren ankommenden Schallsignale zentralnervös vergleichend zu verarbeiten (Kießling et al., 2008). Diese binaurale Interaktion ermöglicht z.B. eine bessere Lokalisation eines Schallsignals im Raum.

Befindet sich eine Schallquelle eher auf der rechten Seite des Kopfes des Hörers, erreicht das von ihr ausgehende Signal das rechte Ohr früher als das linke Ohr. Es entsteht eine Zeitverzögerung zwischen den beiden Ohren, die mit dem Begriff „interaurale Laufzeitdifferenz“ (ITD für engl. interaural time difference) beschrieben wird. Weiterhin erreicht das Signal das rechte, der Schallquelle zugewandte, Ohr mit einem höheren Pegel als das linke, von der Schallquelle abgewandte, Ohr. Dies wird mit dem Begriff des „interauralen Pegel- oder Intensitätsunterschiedes“ (ILD für engl. interaural level difference) beschrieben. Die Auswertung der Laufzeit- und Intensitätsdifferenzen ist frequenzabhängig. Intensitätsunterschiede kann das Gehör nur nutzen, wenn die Wellenlänge des Schalls im

## Kapitel 2. Das auditorische System

Vergleich zum Kopf klein ist. Dies trifft für Frequenzen ab 2 kHz zu. Der Kopf stellt ein Hindernis für die Schallwelle dar, so dass auf der gegenüberliegenden Kopfseite ein Schallschatten entsteht. Bei tieferen Frequenzen und somit langen Wellenlängen stellt der Kopf kein Hindernis dar (Ulrich, 2006), so dass kaum Intensitätsunterschiede wahrgenommen werden. Bei tiefen Frequenzen reagiert das binaurale Hörsystem empfindlich auf Zeitdifferenzen. Weiterhin entsteht durch die Schalleinfallrichtung eine spektrale Veränderung z.B. durch den Kopfschatteneffekt und den Eigenschaften der Ohrmuschel (s. Abbildung 2.4).

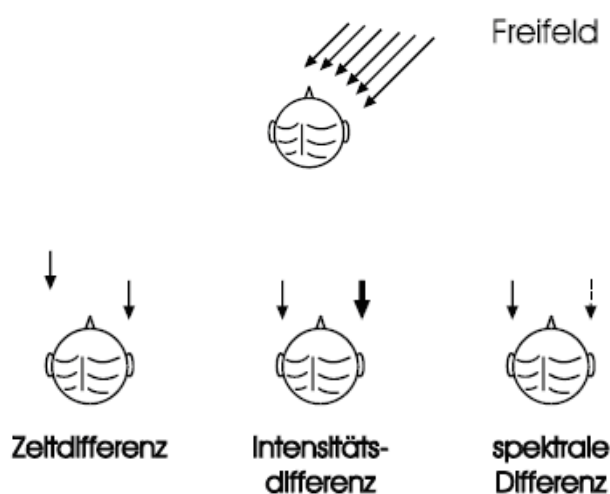


Abbildung 2.4: Physikalische Unterschiede des Schallsignals bei Erreichen beider Ohren bei Schalleinfall von rechts; Links: Zeitdifferenz; Mitte: Intensitätsdifferenz; Rechts: spektrale Differenz (Quelle: Kießling et al., 2008: 28)

Durch die Unterschiede an beiden Ohren bei seitlichem Schalleinfall können selbst kleinste Abweichungen der Schallquelle vom  $0^\circ$  Azimut-Winkel (direkt vor dem Hörer) ausgewertet werden. Die vielfältigen Informationen sind teilweise redundant, so dass selbst in schwierigen Hörsituationen (Hall, Echo, kurze Geräusche), wenn auch eingeschränkter, eine Schallortung möglich ist.

Bei der Schallortung eines Signals, das sich genau vor ( $0^\circ$ ) oder hinter ( $180^\circ$ ) dem Hörer befindet, entstehen weder Laufzeit- noch Intensitätsdifferenzen. Hierbei kommen der Ohrmuschel als auch dem Gehörgang als richtungsabhängige Filter große Bedeutung zu. Durch die Eigenschaften des äußeren Ohres ist eine gewisse Lokalisationsfähigkeit auch unter monauralen Hörbedingungen möglich.

Das beidohrige Hören hat auch einen Einfluss auf die Lautstärkeempfindung. Obwohl sich das physikalische Schallsignal im Pegel nicht verändert, wird es mit zwei Ohren lauter wahrgenommen als mit einem Ohr. Dieses Phänomen ist unter dem Begriff „binaurale Lauthheitssummation“ bekannt. Auf die Hörschwelle bezogen, ergibt sich bei Normalhörenden

## **Kapitel 2. Das auditorische System**

---

ein um 3 dB besserer Wert bei binauralem im Vergleich zum monauralen Hören. Im überschwelligen Bereich kann sich dieser Zugewinn auf 6 - 10 dB belaufen (Kießling et al., 2008).

Normalhörende profitieren von beidohrigem Hören besonders in Kommunikationssituationen mit zusätzlichen Hintergrundgeräuschen (Bronkhorst und Plomp 1988, 1990). Die räumliche Trennung von Nutz- und Störsignal ist ein entscheidender Faktor, durch den bei NH das Sprachverstehen im Störgeräusch verbessert wird (Cherry 1953). Die Verbesserung des Sprachverstehens beruht auf drei Haupteffekten: dem physikalischen Kopfschatteneffekt und den beiden psychoakustisch, binauralen Effekten Squelch und Summation (Durlach und Colburg, 1978).

Der Kopfschatteneffekt tritt auf, wenn Nutz- und Störsignal aus unterschiedlichen Richtungen auf den Hörer treffen. Wenn sich ein Nutzsignal frontal vor dem Hörer und ein Störsignal seitlich vom Hörer befinden, erreicht das Störsignal das zugewandte Ohr schneller und mit einem höheren Pegel als das abgewandte Ohr. Der Kopf stellt ein Hindernis für den ankommenden Schall dar. Das von vorne präsentierte Nutzsignal ist an beiden Ohren gleich. Das im Kopfschatten liegende Ohr hat somit ein lauterer Nutzsignal und ein leiseres Störsignal, erfährt also ein besseres Signal-Rausch Verhältnis (SNR, für engl. signal-to-noise ratio). Untersuchungen zum Kopfschatteneffekt zeigen, dass Normalhörende im Störgeräusch von diesem Effekt profitieren und das Sprachverstehen deutlich verbessert wird (Tillman et al., 1963).

Beim Squelcheffekt sind Nutz- und Störsignal ebenfalls räumlich voneinander getrennt. Durch die Abgleichung der ankommenden Signale an beiden Ohren und somit der Auswertung von ITDs und ILDs, können Nutz- und Störsignal voneinander getrennt werden. Das Störsignal wird unterdrückt und das Nutzsignal besser verstanden. Bei NH konnten Studien zeigen, dass sich die Sprachverständlichkeit aufgrund der binauralen Verarbeitung verbessert, obwohl die Bedingungen für das der Störquelle zugewandte Ohr schwierig sind (Carhart, 1965).

Beim Summationseffekt geht es um die Verbesserung des Sprachverstehens bei beidohrigem Hören im Vergleich zum einohrigen Hören, wenn Nutz- und Störquelle nicht voneinander getrennt sind. Durch Lautheitssummation und redundante Information beider Signale, wird das Sprachverstehen verbessert (Bronkhorst und Plomp, 1990).

Insgesamt entstehen binaurale Vorteile jedoch nur, wenn die Hörer in der Lage sind, die akustischen Informationen beider Ohren zu analysieren und zentralnervös zu fusionieren.

### **2.3 Pathophysiologie des Hörens**

Hörstörungen schränken die Betroffenen in vielen alltäglichen Situationen ein. Personen, die sich jahrelang in der „hörenden“ Welt bewegt haben, verlieren mit einer auftretenden Hörstörung nach und nach ein Mittel zur Kommunikation. Häufig ziehen sie sich bei gesellschaftlichen Anlässen zurück und meiden mehr und mehr soziale Kontakte. Personen, die bereits sehr früh im Leben ihren Hörsinn verloren, haben eventuell noch den Zugang zur Gebärdensprache und können somit ein ihnen geläufiges Kommunikationsmittel verwenden. Nichts desto trotz werden auch sie in der hörenden Welt mit einigen Problemen konfrontiert.

Grundlegend kann zwischen ererbten und erworbenen Schwerhörigkeiten unterschieden werden. Ungefähr 50% der Hörstörungen sind erblich bedingt (Probst, 2008). Sie treten z.B. durch eine gestörte Zellfunktion und -interaktion auf, durch die Proteine fehlerhaft kodiert werden (Lenarz und Boenninghaus, 2012).

Erworbene Hörstörungen werden in die Gruppe der pränatal, perinatal und postnatal auftretenden Hörstörungen eingeteilt. Bei den bereits pränatal, vor der Geburt, auftretenden Hörstörungen zählen Infektionen (Röteln, Zytomegalie etc.), Stoffwechselerkrankungen, Strahlung und gesundheitsschädliche Gewohnheiten der Mutter zu den häufigsten Ursachen (Lenarz und Boenninghaus, 2012). Zu den häufigsten Ursachen einer perinatal, während der Geburt, auftretenden Hörstörung zählen z.B. Untergewicht bei der Geburt und außergewöhnliche Belastungen des Fötus (Geburtstrauma). Zu den postnatalen Hörstörungen zählen alle erworbenen Hörverluste, die nach der Geburt auftreten. Die chronische Mittelohrentzündung stellt hierbei die häufigste Ursache dar. Weitere Ursachen sind die bakterielle Meningitis, Infektions- oder Viruskrankheiten (Probst, 2008).

Zusätzlich kann der Eintritt der Hörschädigung auf den Zeitpunkt des Spracherwerbs bezogen werden. Hierbei unterscheidet man prälingual, perilingual und postlingual auftretende Hörschädigungen.

Zu der prälingual ertaubten Gruppe zählen Personen, bei denen ein Hörverlust vor dem 7. Lebensjahr eingetreten ist. Die bis dahin erworbenen sprachlichen Strukturen und der Wortschatz können wieder verloren gehen. Nach dem 7. Lebensjahr bleibt das akustische Gedächtnis erhalten (Lenarz und Boenninghaus, 2012). Als perilingual ertaubt zählen Personen bis ca. 14 Jahre, die während des Spracherwerbs ihr Gehör verlieren. Sie sind sowohl in ihrer emotionalen als auch in ihrer psychosozialen Entwicklung schwieriger einzuschätzen und liegen zwischen den klarer definierten Gruppen der prä- und postlingual ertaubten Personen (Diller, 2009).

Personen, bei denen die Hörschädigung nach Abschluss des Lautspracheerwerbs plötzlich oder progredient auftritt, zählen zu den postlingual Hörgeschädigten. Darunter fallen

## **Kapitel 2. Das auditorische System**

---

Altersschwerhörige genauso wie junge Erwachsene, die durch Krankheit oder andere traumatische Ereignisse ihre Hörfähigkeit verloren haben.

Grundlegend können periphere von zentralen Hörstörungen unterschieden werden. Bei einer zentralen Hörstörung liegt trotz intaktem peripherem Hörvermögen eine Hörstörung vor. Sie betrifft Hörstörungen oberhalb des Hörnervs. Hierbei handelt es sich um neurologische Störungsbilder, bei denen die auditive Informationsverarbeitung gestört ist (Wolf, 2004).

Bei peripheren Hörstörungen können Außen-, Mittel- und Innenohr sowie der aufsteigende Hörnerv betroffen sein. Diese können in zwei Gruppen eingeteilt werden, die Schalleitungsschwerhörigkeit und die Schallempfindungsschwerhörigkeit.

Die Ursachen für eine Schalleitungsschwerhörigkeit liegen in der Schallaufnahme oder in der Schallübertragung ans Innenohr und betreffen das Außen- und / oder das Mittelohr. Permanente Schalleitungsschwerhörigkeiten können z.B. durch Missbildungen des Gehörgangs, durch Perforation am Trommelfell oder durch Defekte an der Gehörknöchelchenkette entstehen (Lenarz und Boenninghaus, 2012). Vorübergehend kann eine Schalleitungsschwerhörigkeit auch durch Zerumen (Ohrschmalz) oder Mittelohrentzündungen hervorgerufen werden. Die Schalleitungsschwerhörigkeit geht mit einem Intensitätsverlust einher, wobei meist alle Frequenzbereiche gleichermaßen betroffen sind (Wolf, 2004). Für das Sprachverstehen bedeutet dies, dass keine Qualitätsveränderungen auftreten, Sprache jedoch gedämpft wahrgenommen wird. Kennzeichnend ist, dass die Schallübertragung nur über die Luftleitung gestört ist, die Knochenleitung umgeht das Hindernis und ist unbeeinträchtigt. Durch medikamentöse oder apparative Therapien oder operative Maßnahmen kann im Normalfall die Hörfähigkeit nahezu völlig wiederhergestellt werden (Wolf, 2004).

Bei den Schallempfindungsschwerhörigkeiten (sensorineurale Schwerhörigkeit) unterscheidet man je nach Lokalisation auf der Hörbahn die cochleären von den retrocochleären Schwerhörigkeiten.

Bei cochleären Hörstörungen (Innenohrschwerhörigkeit) liegt die Ursache in der Schädigung der äußeren und / oder inneren Haarzellen. Diese Schädigung kann durch Lärmeinwirkung, hörschädigende Medikamente (z.B. einige Antibiotika) oder Substanzen (z.B. Schwermetalle) und Infektionskrankheiten hervorgerufen werden. Im Erwachsenenalter zählen vor allem der Hörsturz, eine Durchblutungsstörung des Innenohres, und die Altersschwerhörigkeit zu den häufigsten Ursachen eines Hörverlustes. Im Tonaudiogramm zeigt sich, dass Luft- und Knochenleitung gleichermaßen zu höheren Pegeln hin verschoben sind (Lenarz und Boenninghaus, 2012).

Sind die äußeren Haarzellen beschädigt, führt dies zu einem Hörverlust für leise und mittellaute Töne, d.h. die Hörschwelle ist zu höheren Pegeln hin verschoben. Durch den

## Kapitel 2. Das auditorische System

---

Ausfall der äußeren Haarzellen kann ein Hörverlust bis zu 60 dB HL (HL für engl. hearing level) entstehen. Laute Töne werden ebenso wahrgenommen wie von Normalhörenden. Auch die Unbehaglichkeitsschwelle zeigt sich bei ähnlichen Pegeln. Die erhöhte Hörschwelle und die gleichbleibende Unbehaglichkeitsschwelle führen insgesamt zu einer veränderten Lautstärkeempfindung. Die empfundene Lautstärke wächst mit steigendem Schallpegel stärker an als bei NH. Dieses gesteigerte Lautstärkeempfinden wird mit dem Begriff „Rekrutment“ bezeichnet (Kießling et al., 2008). Der Ausfall der äußeren Haarzellen führt weiterhin zu einem verschlechterten Zeitauflösungsvermögen, da schnelle Änderungen in der Lautstärke nicht mehr adäquat erfasst werden können. Auch das Frequenzauflösungsvermögen ist beeinträchtigt, was bei der Sprachwahrnehmung dazu führt, dass ähnliche klingende Laute (z.B. /i:/ vs. /e:/) nicht mehr unterschieden werden können (Steffens, 2009). Allgemein ist durch die verringerte Intensitäts-, Frequenz- und Zeitauflösung auch die binaurale Interaktion eingeschränkt. Unterschiede zwischen den Signalen, die an den beiden Ohren ankommen, werden schlechter detektiert. Dies beeinträchtigt sowohl die Schalllokalisierung als auch die Trennung von Nutz- und Störsignalen und somit das Hören in störschallerfüllter Umgebung (Kießling et al., 2008).

Bei einem Hörverlust über 60 dB HL sind zusätzlich zu den äußeren Haarzellen die inneren Haarzellen betroffen. Die Menge der übertragenen Information wird minimiert, da bei Ausfall der inneren Haarzellen keine Möglichkeit mehr besteht, aus den Schallschwingungen Nervensignale zu erzeugen. Bei Ausfall sowohl der äußeren als auch der inneren Haarzellen ist das Hören nicht nur undeutlich, sondern die Informationsübertragung auch lückenhaft. Weiterhin besteht die Gefahr, dass an den Stellen, an denen die inneren Haarzellen ausgefallen sind, auch die Hörnervenfasern verkümmern (Steffens, 2009).

Meistens ist ein Hörverlust nicht über alle Frequenzen gleich. Besonders häufig sind die hohen Frequenzbereiche betroffen (Hochtonschwerhörigkeit). Da der hörbare Frequenzbereich nur unzureichend übertragen wird, führt dies zu Verzerrungen in der Wahrnehmung. Die Wahrnehmung von Konsonanten, insbesondere von Frikativen, ist hierdurch erschwert. Die Schalleitungs- und Schallempfindungsschwerhörigkeit können auch gemeinsam auftreten und somit eine kombinierte Schwerhörigkeit bilden.

Hörstörungen werden weiterhin noch nach ihrem Schweregrad eingeteilt. Der Schweregrad ergibt sich aus der tonaudiometrisch gemessenen Hörschwelle. Nach der World Health Organisation (WHO) spricht man von einer leichtgradigen (26 bis 40 dB Hörverlust), mittelgradigen (41 bis 60 dB Hörverlust), hochgradigen (61 bis 80 dB Hörverlust) und an Taubheit grenzenden Schwerhörigkeit (>81 dB Hörverlust). Es handelt sich hierbei um den Durchschnitt der gemessenen Hörschwelle in dB HL über die Prüffrequenzen 500 Hz, 1 kHz, 2 kHz und 4 kHz.

## **Kapitel 2. Das auditorische System**

---

Je stärker die Schwerhörigkeit, desto schwieriger ist es, diese mit Hörgeräten auszugleichen. Kann das Hörvermögen durch akustische Verstärkung, wie sie durch ein Hörgerät geliefert wird, nicht mehr wiederhergestellt werden, wird die Versorgung mit einem CI in Betracht gezogen (ADANO, 2012).

Für retrocochleäre Hörstörungen, bei denen der Hörnerv z.B. aufgrund eines Tumors nicht mehr intakt ist und somit die neurale Reizübertragung gestört ist (Wolf, 2004), ist eine Versorgung mit einem CI keine Option. Hier können durch ein Hirnstammimplantat (ABI, engl. auditory brainstem implant), bei dem eine Elektrode Strukturen oberhalb des Hörnervs elektrisch reizt, Hörempfindungen ermöglicht werden (ADANO, 2012).



## **KAPITEL 3**

### **Das Cochlea-Implantat**

Die Funktion eines Cochlea-Implantates (CI) besteht in der Überbrückung der Weiterleitung der Schallsignale, wie sie sonst durch die inneren Haarzellen vorgenommen werden würde (Wilson und Dorman, 2008). Das CI ist somit eine Prothese, die zur Rehabilitation von Innenohrschwerhörigkeit eingesetzt werden kann. Wichtigste Voraussetzung für die Indikation eines CIs ist ein intakter Hörnerv. Dieser wird direkt stimuliert, um so eine Hörwahrnehmung zu erzeugen (Lennarz und Boenninghaus, 2012).

#### **3.1 Indikation**

Nach den AWMF-Leitlinien der Arbeitsgemeinschaft deutschsprachiger Audiologen und Neurootologen und Ototologen (ADANO) zur „Cochlea-Implantat Versorgung und zentral-auditorischer Implantate“ vom Mai 2012 kommen für die Versorgung mit einem CI Patienten in Frage, die an einer hochgradigen Schallempfindungsschwerhörigkeit leiden.

Zuerst sollte eine Reihe von Voruntersuchungen stattfinden wie die Ermittlung des Allgemeinzustandes des Patienten und besonders eine HNO-spezifische Anamnese. Hierzu gehören subjektive und objektive Hörprüfungen, die Überprüfung und Optimierung von Hörgeräten, eine neuroradiologische Diagnostik (Felsenbein-CT) und die Magnetresonanztomographie. Durch die Voruntersuchungen sollte sichergestellt sein, dass sowohl der Hörnerv als auch die Hörbahn intakt sind. Neben den audiologischen und medizinischen Voraussetzungen sollten ergänzend pädagogische und psychologische Untersuchungen bzgl. der Kommunikationsfähigkeit, Erwartungshaltung etc. stattfinden. Zusätzlich sollten die Patienten hinsichtlich des Verlaufs einer CI-Versorgung, der unterschiedlichen CI-Systeme, der Rehabilitation und der Erfolgsaussichten aufgeklärt und beraten werden. Erst unter Berücksichtigung aller Befunde wird die Indikation durch den Operateur gestellt.

Das CI sollte zu einem besseren Hören und Sprachverstehen führen, als es mit Hörgeräten zu erwarten wäre. Bei Sprachverständlichkeitstest mit dem „Freiburger Einsilbertest“ ist die

Indikation zum CI gegeben, wenn trotz optimaler Hörgeräteversorgung weniger als 50% der Wörter bei einem Pegel von 70 dB SPL verstanden werden (Helms 1999; Laszig et al., 2004a).

Grundsätzlich ist die Indikation für postlingual (nach dem Spracherwerb) ertaubte und resthörige Kinder, Jugendliche und Erwachsene gegeben. Bei prälingual (vor dem Spracherwerb) ertaubten Erwachsenen ist die Indikation in ausgewählten Fällen gegeben. In den überarbeiteten Leitlinien (bereits in 2009) wird die Indikation bei prälingual ertaubten sowie perilingual (während des Spracherwerbs) ertaubten und resthörigen Kindern empfohlen. Die Versorgung sollte frühzeitig innerhalb der ersten Lebensjahre stattfinden. Eine vorangegangene Beobachtungsphase und eine Hörgeräteversorgung ist jedoch Voraussetzung. Eine Besonderheit stellt der Hörverlust nach bakterieller Meningitis dar. Hierbei kann es zu einer Verknöcherung der Cochlea kommen, die eine intracochleäre Implantation erschweren oder unmöglich machen kann. Eine sofortige audiologische und neuroradiologische Diagnostik ist einzuleiten, die eventuell ohne vorherige Hörgeräteversorgung stattfinden muss (Lenarz et al., 1996).

Hinsichtlich der Versorgung eines oder beider Ohren wurden die Leitlinien ebenfalls überarbeitet. Lange wurde nur die unilaterale CI-Versorgung durchgeführt. Heutzutage ist bei beidseitig gegebener Indikation auch eine beidseitige Implantation durchzuführen (Laszig et al., 2004b). Beidseitig mit CI versorgte Patienten erreichen ein verbessertes Sprachverstehen sowohl in Ruhe als auch im Störgeräusch und verfügen über ein, wenn auch eingeschränktes, Richtungshören (Müller-Deile, 2009). Auch die Höranstrengung wird verringert (Hessel und Hey, 2004).

Mittlerweile kann auch eine einseitige Hörstörung, mit oder ohne zusätzlich begleitendem Tinnitus, welche zu einer erheblichen Einschränkung der Lebensqualität führt, mit einem CI versorgt werden (van de Heyning et al., 2008; Arndt et al., 2010).

Zu den absoluten Kontraindikationen zählen: Eine fehlende Cochlea oder ein fehlender Hörnerv, Taubheit mit Störungen im Bereich der zentralen Hörbahn und eine fehlende Rehabilitationsfähigkeit bzw. eine nicht sichergestellte postoperative Nachsorge.

### **3.2 Aufbau und Funktionsweise eines Cochlea-Implantates**

Das CI besteht aus zwei Hauptkomponenten. Zum einen aus dem Implantat mit Elektrodenträger, zum anderen aus dem extern getragenen Sprachprozessor.

Das Implantat wird während einer mikrochirurgischen Operation unter die Haut implantiert. Es besteht aus der Empfängerspule mit Magnet, der Dekodier- und Stimulatorelektronik, dem Elektrodenträger und einer Referenzelektrode. Am extern, heutzutage meistens hinter dem Ohr, getragenen Sprachprozessor befindet sich die Energieversorgung (Batterie- oder

## Kapitel 3. Das Cochlea-Implantat

Akkufach). Weiterhin verfügt der Sprachprozessor über ein oder mehrere Mikrofone, ein Übertragungskabel und eine Sendespule mit Magnet (s. Abbildung 3.1).

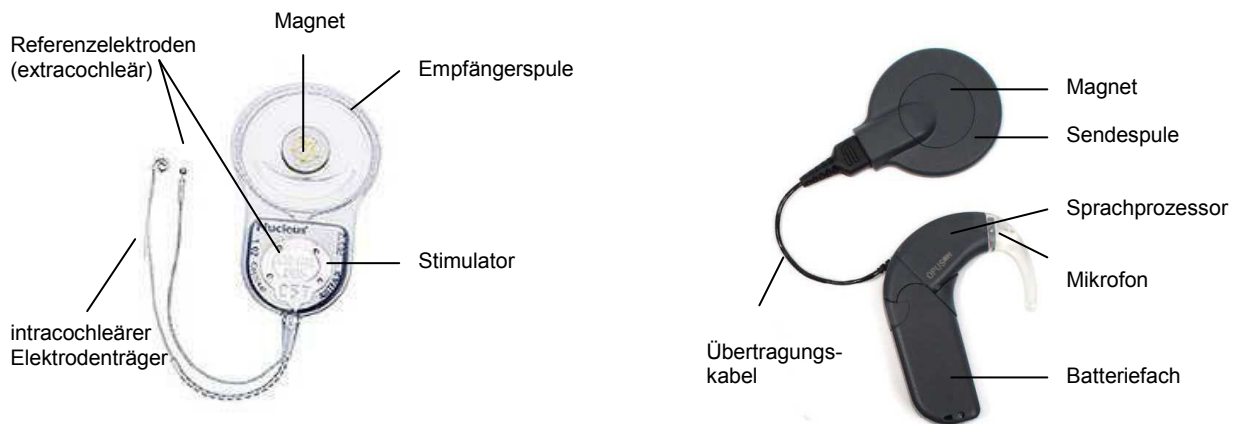


Abbildung 3.1: Interne (Quelle: Cochlear) und externe (Quelle: Med-EI) Komponenten eines Cochlea-Implantat-Systems - modifizierte Darstellung

Aus Abbildung 3.2 wird ersichtlich, wie die Komponenten eines CIs miteinander in Verbindung stehen.

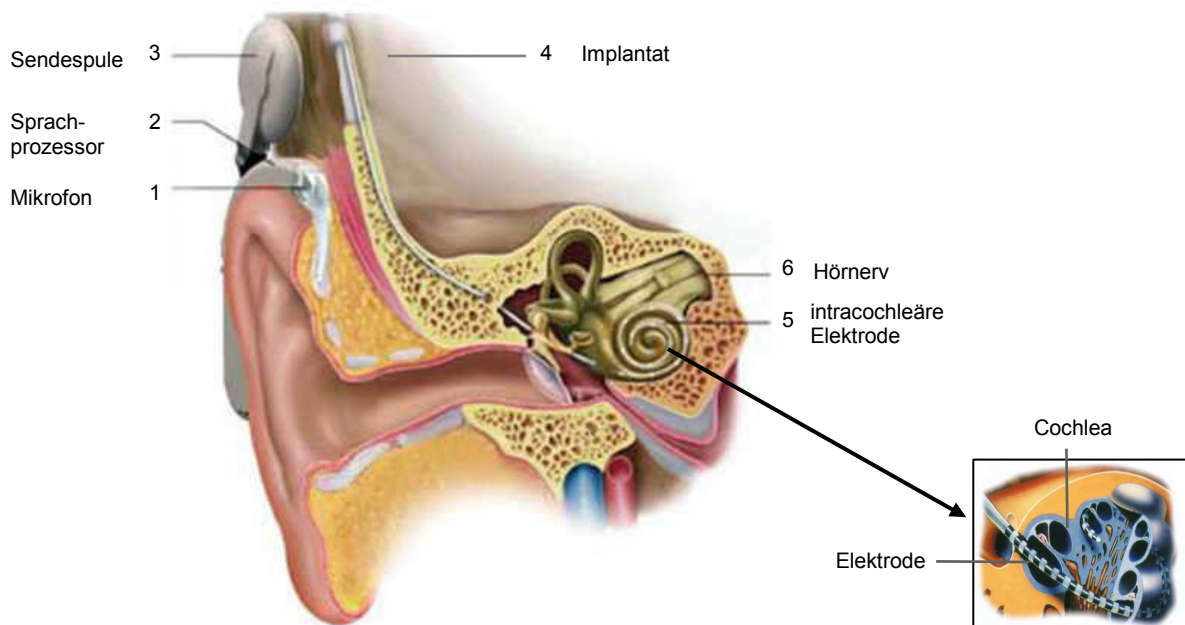


Abbildung 3.2: Schematische Darstellung der Komponenten eines Cochlea-Implantates (Quelle: Med-EI)

Elektrodenlage in der Cochlea (Quelle: Cochlear)

Je nach Hersteller nehmen ein oder mehrere Mikrofone (1), die in den meisten Fällen am Sprachprozessor (2) angebracht sind, die akustischen Schallsignale aus der Umgebung auf. Diese Signale werden in dem Sprachprozessor in elektrische Signale umgewandelt. Eine Sendespule (3), die magnetisch mit dem Implantat verbunden ist, sendet die Signale induktiv durch die Haut (transkutan) an die Empfängerspule im Implantat (4). Mittels der induktiven

### Kapitel 3. Das Cochlea-Implantat

Übertragung wird gleichzeitig die Energieversorgung des Implantates sichergestellt, da dieses über keine eigenständige Energiequelle verfügt. Das Implantat dekodiert die Signale und leitet sie an die Elektrode (5) in der Hörschnecke weiter. Über die elektrischen Impulse der intracochleären Elektrodenkontakte auf der Elektrode (bzw. dem Elektrodenträger) werden die Hörnervenfasern (6) gereizt. Aufgrund des erzeugten elektrischen Feldes ist der Hörnerv in der Lage, Aktionspotentiale zu bilden, so dass das Informationsmuster nach Zentral weitergeleitet wird und im auditorischen Kortex einen Höreindruck erzeugt (Müller-Deile, 2008).

Der Elektrodenträger wird in der Scala tympani in den unteren 1 ½ Windungen der Cochlea platziert (Wilson und Dorman, 2008). Die verfügbaren Elektrodenträger haben eine Länge von 15 mm - 31 mm. Sie werden je nach Hörverlust und Form der Cochlea unterschiedlich tief inseriert. Die Insertion erfolgt entweder über eine Cochleostomie oder über das runde Fenster.

Die zum Zeitpunkt dieser Arbeit verwendeten Systeme sind Mehrkanalsysteme, d.h. es liegen mehrere Elektrodenkontakte innerhalb der Cochlea. Je nach Hersteller befinden sich 12 - 22 Einzelelektroden auf dem Elektrodenträger (Müller-Deile, 2009). Übertragen werden üblicherweise Frequenzen zwischen ca. 100 Hz und 8,5 kHz. Somit sind die für das Sprachverstehen wichtigen Frequenzen in der Übertragung enthalten.

Um die Ladung durch die Elektrodenkontakte an die Nervenfasern zu übertragen, können verschiedene Elektrodenkonfigurationen gewählt werden. Es wird immer eine Referenzelektrode benötigt. Derzeitig verfügbar sind der monopolare, bipolare und common ground Stimulationsmodus (s. Abbildung 3.3).

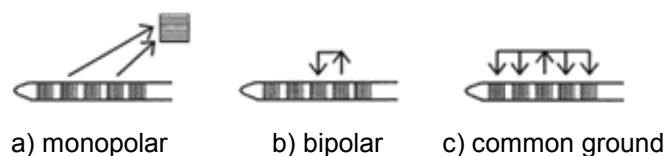


Abbildung 3.3: Darstellung der Stimulationsmodi (a) monopolar, (b) bipolar und (c) common ground (Quelle: McDermott, 2004: 53)

Bei monopolarer Stimulation fließt die Ladung von den aktiven Elektrodenkontakten innerhalb der Cochlea zu einer außerhalb der Cochlea liegenden Referenzelektrode (extracochleär) (s. Abbildung 3.3a). Im bipolaren Stimulationsmodus fließt die Ladung intracochleär zwischen zwei nahe beieinander liegenden aktiven Elektrodenkontakten auf dem Elektrodenträger (s. Abbildung 3.3b). Im Modus „common ground“ fließt die Ladung intracochleär von einem Elektrodenkontakt zu den übrigen Elektrodenkontakten (s. Abbildung 3.3c) (McDermott, 2004).

## Kapitel 3. Das Cochlea-Implantat

Durch die Nutzung mehrerer Elektrodenkontakte können selektiv unterschiedliche Neuronengruppen des Hörnervs stimuliert werden. Allerdings ist es kaum möglich, unerwünschte interferierende elektrische Felder bei eng nebeneinander liegenden Elektrodenkontakten einzuschränken. Zum einen ist der Elektrodenträger in der Scala tympani von leitender Flüssigkeit (Perilyphme) umgeben, zum anderen ist der Elektrodenträger relativ weit vom eigentlich zu stimulierenden Gewebe, den Spiralganglien im Modiolus, entfernt (Wilson und Dorman, 2008). Es besteht die Schwierigkeit, die Ladung, die von einer aktiven Elektrode an Nervenfasern nahe dieser Elektrode abgegeben wird, zu begrenzen. So werden benachbarte Nervenfasern stimuliert, wodurch die Anzahl der effektiv nutzbaren getrennten Elektroden begrenzt ist (Hoth, 1999).

### 3.3 Kodierung akustischer Signale mit dem Cochlea-Implantat

Mithilfe der sog. Kodierungsstrategie werden die akustischen Signale in elektrische Pulse übersetzt. Das Ziel dabei ist, die natürliche Signalanalyse des Ohres nachzuempfinden und durch das elektrische Stimulationsmuster eine vergleichbare Wahrnehmung, wie die eines Normalhörenden, hervorzurufen (Hoth, 1999).

Stimuliert wird meist mit biphasischen ladungsausgeglichenen Pulsen, mit positiver und negativer Phase, um den Elektrolythaushalt der Cochlea nicht zu stören und eine Gewebeschädigung zu verhindern (s. Abbildung 3.4).

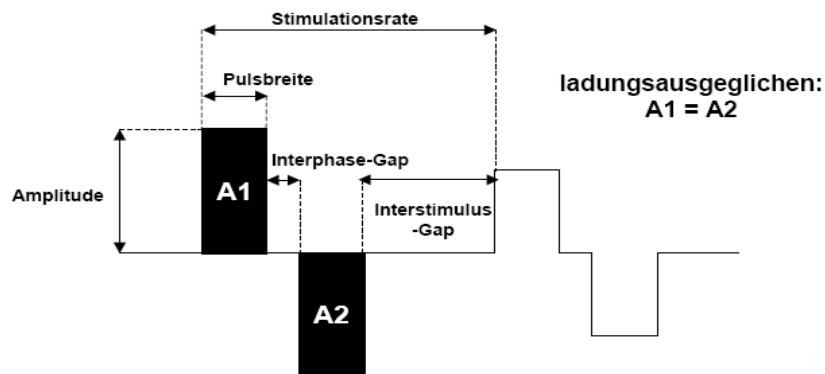


Abbildung 3.4: Biphasisch ladungsausgeglichener Puls

Typische Werte bei einem Nucleus Implantat (Cochlear) Pulsweite:  $25\mu\text{s}$  -  $400\mu\text{s}$ , Pulsamplitude:  $100\mu\text{A}$  -  $1800\mu\text{A}$ , Interphase-Gap:  $8\mu\text{s}$ , Interstimulus-Gap:  $> 360\mu\text{s}$  (Quelle: Hessel, 2004)

Die Parameter der elektrischen Pulse sind die Pulsamplitude (Stromstärke/Ladung) gemessen in Mikroampere ( $\mu\text{A}$ ), die Pulsbreite gemessen in Mikrosekunden ( $\mu\text{s}$ ) und die Puls- oder Stimulationsrate gemessen in Pulsen pro Sekunde (pps) (Hessel, 2004).

Informationen über das Frequenzspektrum erhält man durch die Abbildung der Pulse auf die Elektrodenorte. Unterschiedliche Orte der Stimulation führen zu unterschiedlicher Tonhöhenwahrnehmung. Auf Grundlage der Tonotopie in der Cochlea werden tiefe Frequenzen auf Elektroden nahe des Apex und hohe Frequenzen auf Elektroden nahe der Basis abgebildet. Somit wird die Tonhöhenempfindung einerseits durch die tonotope Frequenzzuordnung erreicht, andererseits kann die Pulsrate ebenfalls die Tonhöhenempfindung beeinflussen (Hoth, 1999).

Je weiter die aktive Elektrode und die Referenzelektrode (extracochleär) auseinander liegen (z.B. im monopolaren Modus), desto ausgedehnter ist die elektrische Feldausbreitung. Der Vorteil hierbei ist, dass eine Stimulation mit niedrigen Reizstärken möglich ist. Je näher die aktive Elektrode und die Referenzelektrode beieinander liegen (z.B. im common ground Modus), desto besser können die Tonotopie und die verfügbaren Elektroden ausgenutzt werden (Hoth, 1999).

Die Zeitstruktur des Signals wird über die Puls- bzw. Stimulationsrate übertragen. Hierbei handelt es sich um die Geschwindigkeit, mit der elektrische Pulse an die Elektroden übertragen werden und die Hörnervenfasern stimulieren. Die zeitliche Auflösung ist durch die Dauer und Frequenz der Pulsfolge begrenzt (Hoth und Müller-Deile, 2009).

Die Intensität eines Signals wird im CI durch die Pulsbreite und die Pulsamplitude, d.h. die abgegebene Stromstärke, kodiert. Bei Erhöhung der Pulsbreite und / oder der Pulsamplitude steigt die Feuerungsrate der Nerven, so dass die Lautheitsempfindung ebenfalls ansteigt. Hierbei ist jedoch darauf zu achten, dass es bereits bei Signalen mit schwacher Intensität zu einer Überlagerung elektrischer Felder kommt. Durch die zunehmend breiteren elektrischen Felder bei Signalen mit hoher Intensität, kommt es zu einer verschlechterten Ortsauflösung, was die Fähigkeit Tonhöhen zu unterscheiden, vermindert. Dadurch ist die sinnvolle Anzahl intracochleärer Elektrodenkontakte begrenzt (Hoth, 1999). Außerdem wird durch die Verbreiterung der Pulse die zeitliche Auflösung verschlechtert (Hessel, 2004).

Wie die Information dieser elektrischen Reize wahrgenommen wird, ist individuell sehr unterschiedlich. Deswegen werden für jede Elektrode Reizstärken ermittelt, die zu definierten subjektiven Empfindungen führen. Zum einen wird eine untere Schwelle ermittelt, bei der ein Reiz gerade eben eine Empfindung auslöst (threshold level = Hörschwelle), und eine obere Schwelle, bei der ein Reiz zu einer angenehm lauten Empfindung führt (comfort level = Behaglichkeitsschwelle). In diesem Dynamikbereich, zwischen elektrischer Hör- und Behaglichkeitsschwelle, bewegen sich die vom Implantat erzeugten Reize. Die Amplitude des akustischen Eingangssignals wird auf den elektrischen Dynamikbereich jeder Elektrode mithilfe einer Wachstumsfunktion („map law“) zwischen Hör- und Behaglichkeitsschwelle abgebildet. Verstärkungs- und Kompressionsmechanismen sorgen dafür, dass leise und laute Signalpegel unterschiedlich übertragen werden. Bei leisen Pegeln ist der Anstieg der

### Kapitel 3. Das Cochlea-Implantat

Wachstumsfunktion steiler, wohingegen laute Pegel abgedämpft werden. Nach der individuellen Einstellung der Schwellen mithilfe der elektrischen Einzelreize, kann der Sprachprozessor aktiviert werden. Die Einstellungen können jederzeit verändert und optimiert werden, so dass der Patient eine möglichst gute Einstellung erhält, mit der Sprache verstanden werden kann (Hoth und Müller-Deile, 2009).

In den von Clark 1977 entwickelten Mehrkanalsystemen übertrugen anfängliche Strategien, wie z.B. die „Feature Extraction Strategy“, wichtige sprachliche Merkmale wie  $F_0$  und  $F_n$  (McDermott, 2004). Bei den heutigen Systemen wird über den gesamten Frequenzbereich hauptsächlich die Information der Einhüllenden übertragen, was die örtliche Kodierung ermöglicht. Eine echte zeitliche Kodierung, anhand der Analyse der Feinstruktur und der damit verbundenen Phasenkopplung bei höheren Frequenzen, findet nicht statt (Med-EI, 2011).

Die grundsätzlichen Mechanismen einer Kodierungsstrategie können anhand der von Wilson et al. (1991) entwickelten „continuous interleaved sampling“ (CIS) Strategie beschrieben werden (s. Abbildung 3.5).

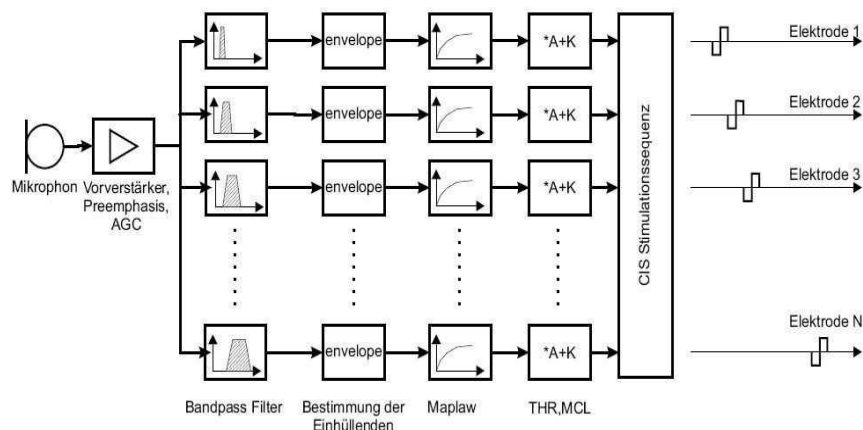


Abbildung 3.5: Blockdiagramm der CIS-Kodierungsstrategie (Quelle: Mitterbacher 2004: 8)

Das durch das Mikrofon aufgenommene akustische Signal wird vorverstärkt und in mehrere Frequenzbänder gefiltert. Jeder Bandpassfilter (Kanal) wird einem Elektrodenkontakt zugeordnet. Die Amplituden der Einhüllenden aller gefilterten Signale werden bestimmt und modulieren die elektrischen Pulse, die an die jeweiligen Elektrodenkontakte weitergegeben werden. Durch Kompression werden die eingehenden, akustisch im Pegel hoch variierenden, Signale auf den engen elektrischen Dynamikbereich im CI abgebildet (Moore, 2003) und somit an den hörbaren Bereich des Patienten angepasst (Mitterbacher, 2004). Die Pulse werden nacheinander an die entsprechenden Elektrodenkontakte weitergegeben, um das bereits beschriebene Problem der Ausbreitung der elektrischen Felder zu verringern. Dies würde bei simultaner Stimulation der

Elektrodenkontakte in höherem Maße produziert werden (Wilson und Dorman, 2008). Außerdem würden sich bei simultaner Stimulation die Nervenfasern immer zur gleichen Zeit in der Erholungsphase (Refraktärzeit) befinden und auf eine erneute Stimulation nicht sofort wieder ansprechen (Müller-Deile, 2009).

Andere Strategien arbeiten nach dem „n aus m“-Prinzip wie z.B. die „spectral peak, SPEAK“ und die „advanced combination encoder, ACE“ Strategie. Grundsätzlich arbeiten die „n aus m“-Strategien ähnlich wie die CIS-Strategie, mit der Ausnahme, dass nur bestimmte Kanäle die Elektrodenkontakte stimulieren. In jedem Stimulationszyklus werden die Amplituden der Einhüllenden in den Kanälen miteinander verglichen. Nur die Signalanteile mit den n-höchsten Amplituden werden aus den m-Bandpassfiltern an die intracochleären Elektrodenkontakte übertragen. Die Anzahl der Kanäle entspricht nicht der Anzahl der Elektrodenkontakte, die später stimuliert werden. Die Anzahl n ist bei der ACE-Strategie festgelegt, wohingegen die Anzahl bei der SPEAK-Strategie variieren kann. Ein weiterer Unterschied zwischen ACE und SPEAK ist die Stimulationsrate. Während bei der ACE-Strategie mit nahezu 1000 Pulsen pro Sekunde pro Elektrode stimuliert wird, ist die Stimulationsrate bei der SPEAK-Strategie mit ca. 250 Pulsen pro Sekunde sehr viel langsamer.

Da bei den „n aus m“-Strategien immer nur eine bestimmte Anzahl von Elektroden die Hörnervenfasern stimuliert, sollte die gegenseitige Beeinflussung von Elektroden und die Anzahl der aktivierten Nervenfasern in der Cochlea minimiert bzw. präzisiert werden (Wilson und Dorman, 2008).

Ein neuerer Ansatz, entwickelt von Zierhofer (1999), kodiert die Feinstruktur in den tieferen Frequenzbereichen eines Signals (fine structure processing, FSP). Zusätzlich zur Einhüllenden wird die Feinstruktur durch Stimulationsimpulse übertragen, die bei den Nulldurchgängen des bereits gefilterten Signals ausgelöst werden (Mitterbacher, 2004). Diese Auslösung der Pulse findet maximal in den vier apikalsten Elektroden statt, während die danach folgenden höheren Elektroden wiederum durch eine konstante Pulsrate die Information der Einhüllenden abtasten. Der Übertragungsbereich beginnt in den tiefen Frequenzen bei etwa 70 Hz. Durch diese Neuerung soll insbesondere das Sprachverstehen im Störgeräusch und die Musikwahrnehmung verbessert werden (Smith et al., 2002).

Im Sprachverstehen führen die zurzeit genutzten Kodierungsstrategien jedoch allesamt zu vergleichbaren Ergebnissen (Zeng, 2011).

### **3.4 Versorgungsarten**

Neben der unilateralen CI-Versorgung gehören die bimodale Versorgung (CI und HG) und die bilaterale Implantation (zwei CIs) zur Regelversorgung. Bimodal und bilateral versorgte



Patienten gehören damit zur größten Patientengruppe und werden im folgenden Unterkapitel detaillierter beschrieben.

Aufgrund neuerer Konzepte können auch Patienten mit einem CI versorgt werden, die auf diesem Ohr noch über Restgehör verfügen. Hierbei handelt es sich um die elektroakustische Stimulation (EAS) (v. Ilberg, 1999) oder Hybridstimulation, bei der ein Ohr sowohl mit einem HG als auch mit einem CI versorgt ist. Hiervon profitieren vor allem Patienten mit einem Hörverlust im Hochtonbereich und ausreichendem Restgehör im Tieftonbereich. Ein mittelgradiger Hörverlust ( $\leq 65$  dB HL von 250 Hz bis 750 Hz) wird durch einen Hörgeräteanteil im Sprachprozessor verstärkt, während der hochgradige Hörverlust ( $> 75$  dB HL bei 1,5 kHz bis 8 kHz) über elektrische Stimulation durch das CI ausgeglichen wird (Maurer, 2009). Im Vergleich zur normalen CI-Versorgung wird ein kürzerer Elektrodenträger verwendet (16 bis 24 mm, je nach Hersteller), um den Erhalt des Restgehörs bei der Operation zu sichern. Die Elektrodenanzahl ist die gleiche wie bei einer Standardelektrode, jedoch ist der Abstand benachbarter Kontakte verringert. Diese Patientengruppe profitiert von einer besseren Hörqualität und von einem verbesserten Sprachverstehen im Störgeräusch (Baumann und Helbig, 2009).

Eine weitere neue Versorgungsoption ist die Implantation eines CIs bei einseitiger Taubheit. Einige Patienten, die auf einem Ohr normalhörend sind, können durch das CI auf der hochgradig an Taubheit grenzenden kontralateralen Seite wieder das binaurale Hören zurückgewinnen.

Dies hilft besonders beim Richtungshören und beim Verstehen im Störgeräusch (Jacob, 2011). Außerdem wirkt sich die CI-Versorgung positiv auf die Lebensqualität aus. Im Falle von zusätzlichem Tinnitus kann eine CI-Versorgung zu einer Reduktion des Ohrgeräusches führen (Arndt et al, 2010).

### **3.4.1 Bimodale und bilaterale Versorgung**

Häufig wurde vor der CI-Implantation bereits versucht den Hörverlust mit HGs auszugleichen. Genügte diese Art der Versorgung nicht, um hinreichend Sprache zu verstehen, kann, häufig erst für eine Seite, die CI-Implantation in Betracht gezogen werden. Bei dieser Gruppe der Patienten, die mit einem CI und einem HG auf der kontralateralen Seite versorgt sind, spricht man von der „bimodalen Versorgung“. Im Vergleich zur alleinigen Nutzung des CIs können Patienten beim Sprachverstehen in Ruhe und im Störgeräusch von einem zusätzlichen Hörgerät profitieren (Ching et al., 2007). Die bimodale Nutzung wird für resthörige Patienten empfohlen (Ching et al., 2004).

Da bei den meisten Ausprägungen einer Schallempfindungsschwerhörigkeit das Restgehör in den tiefen Frequenzen besser erhalten ist als in den hohen Frequenzen, ist der

Höreindruck mit dem HG meist tiefer als mit dem CI. In der Regel können die unterschiedlichen Klänge beider Ohren und somit die Kombination von elektrischer und akustischer Stimulation zentral zum Hören und Sprachverstehen fusioniert werden (Müller-Deile, 2009). Die Erweiterung um die tiefen Frequenzen wird als angenehm empfunden und trägt zur Natürlichkeit des Klangeindrucks bei. Das Hören der tiefen Frequenzen hat ebenfalls den Vorteil, dass eine bessere Repräsentation von F0 und stimmhafter Anteile in Sprache vorliegt. Somit ist das Sprachverstehen in der bimodalen Kondition meist besser als in der unilateralen CI-Kondition (Zhang et al., 2010).

Es profitieren jedoch nicht alle Patienten gleichermaßen von der bimodalen Versorgung. Die bimodale Hörsituation ist der unilateralen CI-Nutzung dann überlegen, wenn noch ausreichend Restgehör auf dem mit dem HG versorgten Ohr vorhanden ist. Morera et al. (2005) zeigten, dass das Sprachverstehen mit beiden Hörhilfen kaum besser war als mit dem CI alleine, wenn die Patienten kaum Restgehör aufwiesen und nur wenig mit dem HG alleine verstanden.

Den Patienten wird empfohlen, das HG auf der kontralateralen Seite so lange wie möglich zu tragen, um Deprivationsfolgen vorzubeugen. In dem Fall, dass das HG weder einen ersichtlichen Nutzen im Sprachverstehen bringt, noch eine Hilfe für die räumliche Orientierung darstellt, legen und lehnen die Patienten das HG nach kurzer Tragedauer ab (Müller-Deile, 2009). Gegebenenfalls besteht dann die Option, dieses Ohr mit einem zweiten CI zu versorgen.

Die Versorgung mit zwei CIs bezeichnet man als „bilaterale Versorgung“. Ähnlich wie bei der bimodalen Versorgung zeigen bilateral versorgte CI-Träger Verbesserungen im Sprachverstehen sowohl in Ruhe als auch im Störgeräusch (Laszig et al., 2004b; Müller et al., 2002). Auch das Richtungshören und die räumliche Orientierung verbessern sich im Vergleich zur unilateralen CI-Nutzung (Müller-Deile, 2009). Ein zweites CI verändert die Klangwahrnehmung. Patienten beschreiben subjektiv ein volleres und runderes Klangerlebnis (Hessel und Hey, 2004).

Sowohl für bimodal als auch für bilateral versorgte CI-Träger gilt, dass die Höranstrengung abnimmt und konzentriertes Zuhören über eine längere Zeitspanne möglich ist (Hessel und Hey, 2004).

### **3.4.2 Exkurs Hörgeräte**

Da das Probandenkollektiv dieser Arbeit größtenteils bimodal, mit einem CI und einem Luftleitungshörgerät, versorgt war, wird ein kurzer Überblick über den Aufbau und die Funktion von Hörgeräten gegeben.

### Kapitel 3. Das Cochlea-Implantat

In den Richtlinien des Bundesausschusses der Ärzte und Krankenkassen (Hilfsmittel-Richtlinien) sind die folgenden Indikationskriterien für eine Hörgeräteversorgung festgesetzt: Eine HG-Versorgung sollte erfolgen, und wird von den gesetzlichen Krankenkassen finanziell bezuschusst, wenn in den tonaudiometrischen Messungen in dem Bereich von 500 Hz bis 3 kHz die gemessene Hörschwelle mindestens einer Prüffrequenz 30 dB HL übersteigt. In den sprachaudiometrischen Messungen sollten bei Einsatz des Freiburger Einsilbertests ohne Hörhilfe höchstens 80% bei 65 dB SPL verstanden werden. In den Hilfsmittel-Richtlinien spielen auch andere Faktoren wie das Kommunikationsvermögen des Patienten oder die Akzeptanz von Hörgeräten eine große Rolle (Kießling et al., 2008).

Grundlegend nimmt ein HG den Schall aus der Umgebung über ein oder mehrere Mikrofone auf, verstärkt diesen je nach Hörverlust und gibt ihn in der veränderten Form an das Trommelfell ab. Am häufigsten werden heutzutage digitale HGs verwendet. Hierbei wird, wie in Abbildung 3.6 ersichtlich, der akustisch ankommende Schall von einem oder mehreren Mikrofonen (Empfänger) aufgenommen und in eine elektrische Wechselspannung umgewandelt. Dieses elektrische Eingangssignal wird vorverstärkt und mittels eines analog-digital coders (ADC) digitalisiert. Dieses Digitalsignal wird in Frequenzbänder zerlegt und entsprechend den Bedürfnissen des Hörgerätenutzers frequenzspezifisch angepasst, d.h. es wird bzgl. der Lautstärke, der Frequenzverarbeitung etc. modifiziert. Danach wird das Signal in ein analoges Signal, mittels eines digital-analog coder (DAC), zurück transformiert und nachverstärkt. Ein Hörer (Lautsprecher) gibt das so bearbeitete Signal an das Ohr weiter. Eine Batterie sorgt für die benötigte Energie (Kießling et al., 2008).

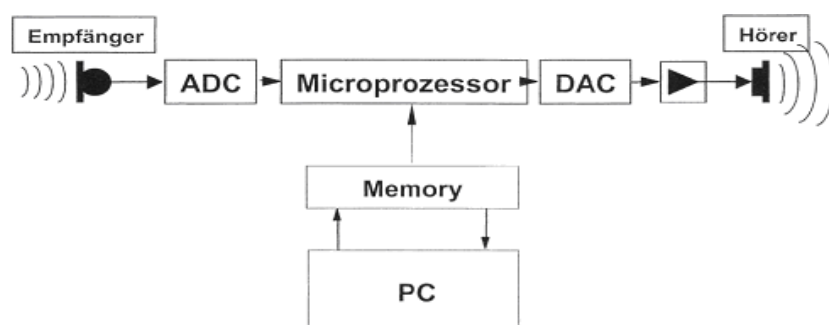


Abbildung 3.6: Schematische Darstellung des Aufbaus eines digitalen Hörgerätes; ADC: Analog-digital coder; DAC: Digital-analog coder (Böhme und Welzl-Müller, 2005: 234).

Bei dem sog. „Hinter-dem-Ohr“ (HdO)-Gerät befinden sich alle Komponenten in einem Gehäuse, das, wie der Name vorgibt, hinter dem Ohr getragen wird. Das verstärkte Signal wird über einen Hörer in einen Schallschlauch abgestrahlt, der den Schall in den Gehörgang bzw. zum Trommelfell leitet. Bei einigen Geräten ist der Hörer direkt in einem Ohrpassstück im Gehörgang angepasst. Weiterhin gibt es „In-dem-Ohr“ (IdO)-Geräte. Diese werden entweder in der Ohrmuschel oder im Gehörgang getragen. Grundsätzlich ist der Ort der

Schallaufnahme bei IdO-Geräten natürlicher als bei HdO-Geräten. Die Nutzung der jeweiligen Geräte ist allerdings abhängig von der Art und Schwere des Hörverlustes. Während HdO-Geräte bei jedem Hörverlustgrad verordnet werden können, sind IdO-Geräte nur für einen leicht- bis mittelgradigen Hörverlust sinnvoll einsetzbar (Böhme und Welzl-Müller, 2005).

Moderne Hörgeräte werden über einen PC durch speziell entwickelte Anpassprogramme vom Akustiker eingestellt. Der Einstellung von Hörgeräten und somit dem Zugewinn zum Sprachverstehen sind jedoch Grenzen gesetzt. Die Funktion der äußeren Haarzellen kann eingeschränkt durch ein HG wiederhergestellt werden. Durch nicht-lineare Verstärkung ist es möglich, die Lautheitsempfindung wieder an die eines NH anzugleichen. Je nach Hörverlust können leise Signale theoretisch so weit verstärkt werden, dass man nahe an die Hörschwelle Normalhörender herankommt. Laute Signale hingegen würden bei nicht-linearer Verstärkung automatisch gedämpft werden (Steffens, 2009). Hinsichtlich des Zeitauflösungsvermögens gibt es bis jetzt noch keine Möglichkeit, diesen Zustand mit HGs zu verbessern. Die Verbesserung der Frequenzauflösung ist eingeschränkt möglich, eine völlige Normalisierung kann jedoch nicht erreicht werden. Es besteht weiterhin eine Verschlechterung des Richtungshörens (Steffens, 2009). Auch Verzerrungen und fehlende Tonbereiche können nicht behoben werden (Wolf, 2004).

Bei einem Ausfall der inneren Haarzellen kann deren Funktion nicht mithilfe eines HGs kompensiert werden (Steffens, 2009). Hier könnte das Hörvermögen nur noch durch die Versorgung mit einem CI wiederhergestellt werden.

#### **3.5 Möglichkeiten und Grenzen beim Hören mit dem Cochlea-Implantat**

Die akustischen Größen werden bei der Verarbeitung und Übertragung mit einem CI verändert. Jedoch ist trotz eingeschränkter Frequenzauflösung sowie limitierter Abtast- und Stimulationsrate (Rubinstein, 2004) ein offenes Sprachverstehen in ruhiger Umgebung meist möglich (Skinner, 2003). Allerdings ist das Sprachverstehen von CI-Trägern inter-individuell sehr unterschiedlich. So sind einige CI-Nutzer in der Lage, Sprache im Störgeräusch zu verstehen und zu telefonieren (Anderson et al., 2006), während anderen CI-Nutzern „nur“ das Sprachverstehen in Ruhe mithilfe des Mundbildes des Sprechenden gelingt.

Einige Faktoren beeinflussen die Entwicklung des Hörens mit dem CI. Ein wichtiger Faktor ist der Zeitpunkt der Hörschädigung. Patienten, die postlingual hörgeschädigt sind und mit einem CI versorgt wurden, erlangen in der Regel ein besseres Sprachverstehen mit dem CI als prälingual hörgeschädigte Patienten. Weiterhin spielt die Dauer der Hörschädigung eine große Rolle. So hat eine lange Dauer der Hörschädigung einen negativen Effekt auf die Hör- und Sprachwahrnehmung, da die Hörnervenfasern verkümmern können. Das Tragen von

HGs vor der Implantatversorgung hat einen positiven Effekt auf die Entwicklung der Sprachwahrnehmung. Es verhindert die Deprivation der Hörnervenfasern wenigstens teilweise. Ebenso ist das Alter bei Implantatversorgung wichtig. Besonders bei prälingual hörgeschädigten Patienten zeigt sich ein Unterschied im Sprachverstehen, je nach dem ob sie als Kinder oder Jugendliche oder im Erwachsenenalter versorgt wurden (Loizou, 1998). Immer häufiger steht auch das Alter der Patienten im Fokus der Diskussion. So zeigten einige Studien, dass ältere Patienten (>60 Jahre) weniger von einem CI profitieren als jüngere Patienten, wobei auch jene ältere Patientengruppe einen profitablen Hörerfolg erzielen kann. Einen noch höheren Faktor als die Dauer der Hörschädigung und das Alter der Patienten stellt das Höralter, also die Erfahrung mit dem CI, dar. Je länger das CI getragen wird, desto besser sind die Fähigkeiten, Sprache wahrzunehmen (Blamey et al., 2013).

Hinzu kommen technische Faktoren wie z.B. die Elektrodenlage und die Kodierungsstrategie. Auch die Anzahl der aktiven Elektroden sowie die Bedeutung der Insertionstiefe des Elektrodenträgers stehen zur Diskussion. Häufig ist es nicht möglich, die Bedeutung einzelner Faktoren unabhängig voneinander zu beurteilen, da immer eine Vernetzung mehrerer Faktoren vorliegt (Loizou, 1998). Unabhängig von den patientenspezifischen und technischen Faktoren zeigt sich in nahezu allen Bereichen der Hörwahrnehmung über ein CI, dass postoperatives Hörtraining maßgeblich zum Erfolg des Hörens und im Speziellen des Sprachverstehens beiträgt (Hoth, 1999).

Im Vergleich zu Normalhörenden zeichnet sich jedoch ein schlechteres Sprachverstehen im Störgeräusch, Probleme bei der Tonhöhen- und Musikwahrnehmung und in der Identifizierung und Unterscheidung von Stimmen ab (Zeng, 2011).

### **3.5.1 Sprechererkennung mit CI**

Die eingeschränkte Auflösung akustischer Merkmale im CI und besonders tieffrequenter Information führt dazu, dass stimmliche Charakteristiken wie  $F_0$  und  $F_n$  nur eingeschränkt dazu beitragen können, Sprecher zu erkennen (Vongphoe und Zeng, 2005). Die Verarbeitung und Übertragung von  $F_0$  ist für CI-Träger weder aufgrund der Ortskodierung noch aufgrund der zeitlichen Kodierung verlässlich. Zu wenige Elektrodenkontakte und eine zu große Ausbreitung der Ladung und Diskrepanzen zwischen Eingangssignal und den durch die Elektroden gereizten Hörnervenfasern tragen zu einer ungenauen Ortskodierung und zu einer schlechten spektralen Auflösung für  $F_0$  bei (Rubinstein, 2004). Es besteht die Annahme, dass die  $F_0$ -Information durch zeitliche Kodierung (Periodizität) übertragen wird und dies, vor allem bei großen  $F_0$ -Unterschieden, dem Erkennen des Sprechergeschlechts dient (Meister et al., 2009).

Dies folgerten auch Fu et al. (2005), die Experimente mit CI-Trägern zur Erkennung des Sprechergeschlechts durchführten. Unterschiede von durchschnittlich 10 Hz zwischen den Stimmen konnten kaum genutzt werden, um das Sprechergeschlecht eindeutig zuzuordnen zu können. Bei großen F<sub>0</sub>-Unterschieden von durchschnittlich 100 Hz gelang den CI-Trägern jedoch die Zuordnung des Sprechergeschlechts zu 85%. Auch in einer Studie von Meister et al. (2009) konnte gezeigt werden, dass Sprecher unterschiedlichen Geschlechts erkannt wurden und die Ergebnisse mit denen von NH vergleichbar waren. Kovacic und Balaban (2009) fanden aufgrund ihres Versuchsaufbaus heraus, dass ihre CI-Probanden, männliche und weibliche Stimmen voneinander unterscheiden konnten, wenn diese nacheinander präsentiert wurden. Wurde eine Sprecherstimme nur einzeln angeboten, gelang die Identifikation des Sprechergeschlechts nicht. Sie folgerten daraus, dass zusätzlich auch Prozesse wie die Präsentation und Verarbeitung auditorischer Objekte im Gedächtnis eine Rolle spielen.

Geht man über die Erkennung und Unterscheidung des Sprechergeschlechts hinaus und widmet sich der Erkennung und Unterscheidung von Sprechern des gleichen Geschlechts, sind die Probleme von CI-Trägern weitaus größer. Hierbei spielt nicht nur F<sub>0</sub> eine Rolle, sondern auch die Verarbeitung der F<sub>n</sub> in höheren Frequenzbereichen. Die Wahrnehmung von F<sub>n</sub> ist zwar besonders wichtig, aber durch die Verarbeitung im CI eingeschränkt (Vongphoe und Zeng, 2005).

#### **3.5.2 Sprachverstehen im Störgeräusch mit CI**

Das Sprachverstehen im Störgeräusch beruht auf der Auswertung sowohl feiner Amplitudenunterschiede als auch zeitlicher und spektraler Merkmale der Sprache, deren Verarbeitung in CIs jedoch stark eingeschränkt ist. Gründe dafür sind z.B., dass es bei Stimulation des Hörnervs durch die Ausbreitung des elektrischen Feldes zu Kanalinteraktionen kommt (Rubinstein, 2004; Stickney et al., 2006). Weiterhin fehlt die präzise Übertragung der Feinstruktur der Signale, da die CI-Systeme größtenteils die Amplitudeneinhüllende des Signals analysieren, um Sprache zu übertragen (Nie et al., 2006). Wird Feinstruktur nur wenig oder gar nicht wahrgenommen, ist das Sprachverstehen im Störgeräusch erschwert (Qin und Oxenham, 2003).

Die Auswertung und Übertragung der F<sub>0</sub> spielt eine sehr bedeutsame Rolle für das Sprachverstehen mit Hintergrundgeräuschen oder mehreren Sprechern. Die Schwierigkeiten von CI-Trägern, Sprache in komplexen Hörsituationen zu verstehen, wird teilweise damit erklärt, dass F<sub>0</sub> nur schlecht durch das CI übertragen wird. Trotz der verschlechterten F<sub>0</sub>-Wahrnehmung können CI-Träger unterschiedliche auditive Ströme erkennen und voneinander trennen. In Bezug auf die Differenzierung unterschiedlicher, gleichzeitig

präsentierter Vokale, bei denen spektrale Merkmale durch eine CI-Simulation reduziert waren, zeigte sich, dass F0 die Zusammenfassung zu einem auditiven Strom teilweise erlaubt. Für die meisten CI-Träger bestehen aber deutliche Schwierigkeiten (Gaudrain et al., 2008).

Bedingt durch die Erschwernisse, F0 exakt wahrzunehmen, ist das Lückenhören für CI-Träger ebenfalls schwierig. Das Sprachverstehen, besonders im fluktuierenden Störgeräusch, gelingt kaum. Weder zeitliche noch spektrale Lücken in simultanen Schallsignalen können genutzt werden (Nelson et al., 2003). Wie bereits erwähnt basiert die Wahrnehmung bei CI-Trägern auf der Information der Einhüllenden eines Schallsignals. Die Einhüllende bildet jedoch bei konkurrierenden Signalen keine verlässliche Basis mehr, die für das Sprachverstehen genutzt werden kann (Nelson und Jin, 2004).

Anlehnend an eine Studie von Binns und Culling (2007) führten Meister et al. (2011b) eine Studie mit CI-Trägern durch, in der zum einen die Unterscheidungsfähigkeit von F0-Änderungen und zum anderen das Sprachverstehen im Störgeräusch gemessen wurde. Beim Sprachverstehen im Störgeräusch wurde die F0 des Zielsprechers, und damit die Intonationskontur, entweder unverändert oder invertiert dargeboten. In den Ergebnissen zeigte sich, dass das Sprachverstehen der CI-Träger schlechter war, wenn die F0 des Sprechers invertiert war. Zusammenhänge mit der Fähigkeit F0-Unterschiede wahrzunehmen bestanden. Dies zeigt, dass die Wahrnehmung von Intonation und damit von F0, das Sprachverstehen im Störgeräusch von CI-Trägern beeinflusst.

Individuell sind besonders in Störgeräuschsituationen und beim Telefonieren, also in Situationen, in denen das Sprachsignal den Hörer nicht ungestört erreicht, kompensatorische Fähigkeiten wichtig. Normalhörenden helfen das linguistische Wissen und die Erschließung des Kontextes, ungehörte oder verdeckte Sprachanteile vorherzusagen, wodurch das Sprachverstehen im Störgeräusch verbessert wird. Diese Vorhersageprozesse (top-down Prozesse) sind aufgrund eingeschränkter Signalqualität und dadurch verschlechterter bottom-up Prozesse für CI-Träger vermindert (Başkent, 2012).

### **3.5.3 Prosodie- und Musikwahrnehmung mit CI**

Die Intensität, die Quantität und die Intonation bilden den Bereich der linguistischen Prosodie. Diesen Parametern sind die akustischen Größen Amplitude, Dauer und F0 zugeordnet, wobei diese Zuordnung keineswegs eindeutig ist.

### Kapitel 3. Das Cochlea-Implantat

---

Es besteht eine gute Korrespondenz zwischen Quantität und Dauer sowie zwischen Intonation und Sprachgrundfrequenz. Der akustische Parameter Energie spielt jedoch eine untergeordnete Rolle: Betonung und Akzentuierung wirken vor allem über die Dauer und über die Sprachgrundfrequenz. Daneben besteht ein starker Einfluss der Intensität auf den Artikulationsaufwand: Betonte Silben werden stets sorgfältiger und präziser artikuliert als unbetonte: daher verlangen sie auch mehr Zeit und werden dementsprechend gelängt (Hess, 2003: 7).

Lange war unklar, welche der oben genannten Parameter für CI-Träger von besonderer Bedeutung sind, um prosodische Merkmale wahrzunehmen. Bei der Entwicklung einer deutschsprachigen Prosodie-Testbatterie (Meister et al., 2008) wurden prosodische Merkmale in vier Modulen gezielt untersucht. Es gab Minimalpaare, die sich hinsichtlich ihrer Vokallänge unterschieden. Weiterhin wurde die Unterscheidung von Frage und Aussagesätzen mit gleicher Satzstruktur, aber unterschiedlichem F<sub>0</sub>-Verlauf getestet. Ähnlich war auch der Test bzgl. unterschiedlicher Satzakzente, bei dem sich ebenfalls nur der F<sub>0</sub>-Verlauf änderte. Das 4. Modul bestand aus der Unterscheidung des Sprechergeschlechts. Insgesamt zeigte sich, dass CI-Träger vor allem die Signaldauer ausnutzen konnten. Änderungen der F<sub>0</sub> reichten weniger aus, um die linguistische Funktion prosodischer Merkmale zu entschlüsseln. In einer weiterführenden Studie (Meister et al., 2011a), in der vor allem der Satzakzent untersucht wurde, wurden Stimuli künstlich manipuliert. Dadurch konnten die Merkmale der Amplitude, der Dauer und F<sub>0</sub> getrennt voneinander verändert werden. CI-Träger und NH profitierten gleichermaßen von Änderungen in der Dauer. Merkmale der Amplitude sowie Merkmale der F<sub>0</sub> konnten von den CI-Trägern jedoch nicht genutzt werden, um unterschiedliche Betonungen zu differenzieren. Dies ist für tonale Sprachen von großer Bedeutung, da es in Sprachen wie dem Chinesischen Wörter gibt, die sich durch den unterschiedlichen F<sub>0</sub>-Verlauf in ihrer Bedeutung unterscheiden (Hess, 2003). Eng mit der Unterscheidung von F<sub>0</sub>-Konturen bzw. dadurch wahrgenommener Tonhöhenkonturen ist die Melodieerkennung bei der Musikwahrnehmung verbunden. Die Verbesserung der Musikwahrnehmung ist neben der Verbesserung der Sprachwahrnehmung ein weiterer Bereich, um die Lebensqualität von CI-Trägern zu erhöhen. Hierbei spielen allerdings nicht nur einzeln zu differenzierende Merkmale eine Rolle, sondern das Zusammenspiel vieler unterschiedlicher Parameter. Die Musikwahrnehmung ist eine hoch komplexe Leistung des auditorischen Systems (McDermott, 2004).

Drei grundlegende Parameter der Musik sind der Rhythmus, die Tonhöhe und die Klangfarbe. Die Wahrnehmung von Rhythmus basiert auf der Wahrnehmung der Länge eines Tons und den Pausen zwischen den Tönen. Hierbei spielt die zeitliche Auflösung und Verarbeitung die wichtigste Rolle. Diese ist bei der Kodierung in einem CI ausreichend umgesetzt. McDermott (2004) beschreibt in einer Zusammenfassung über unterschiedliche Studien zur Musikwahrnehmung von CI-Trägern, dass die Wahrnehmung von Rhythmus



besonders gut gelingt und ähnlich der von NH ist (Kong et al., 2004). Die Unterscheidung komplexer Rhythmen bleibt hingegen schwierig. Hier bleiben die Ergebnisse der CI-Träger hinter den Ergebnissen der NH zurück (Drennan und Rubinstein, 2008).

Die Wahrnehmung von Melodie ist schwierig für Implantatträger. Auch wenn die Melodie eines Liedes CI-Trägern bekannt ist, können die Lieder nicht einwandfrei identifiziert werden. In einigen Studien zeigte sich, dass CI-Träger Melodien erkannten, die ein ausreichend zu unterscheidendes rhythmisches Muster aufwiesen, im Gegensatz zu Melodien, die wenig Rhythmus beinhalteten (Kong et al., 2004). Die eingeschränkte Melodiewahrnehmung ohne die Nutzung rhythmischer Parameter wird durch die eingeschränkte Wahrnehmung der Tonhöhe erklärt. Dies entspricht in etwa den Problemen, wie sie bei der Übertragung und Wahrnehmung der F0 vorkommen.

Durch den Parameter Klangfarbe können Töne voneinander unterschieden werden, die von gleicher Lautstärke, Tonhöhe und Dauer sind. So ist es z.B. möglich, verschiedene Instrumente, die den gleichen Ton spielen, voneinander zu unterscheiden. Normalhörende nutzen Merkmale, die sowohl aus der Information der Einhüllenden als auch aus Information der Feinstruktur abgeleitet werden können. Die verschlechterte Wahrnehmung klanglicher Parameter im Vergleich mit NH (Kong et al., 2011) ergibt sich bei CI-Trägern dadurch, dass sie hauptsächlich Informationen der Einhüllenden nutzen können und nicht die der Feinstruktur.

### **3.5.4 Beidohriges Hören mit Hörhilfen**

Bei postlingual hörgeschädigten Erwachsenen hat sich das binaurale Hören unter „normalen“ Bedingungen entwickelt, so dass besonders diese Patientengruppe auf eine Wiederherstellung des binauralen Systems hoffen kann.

Jedoch besteht bei allen technischen Hörhilfen das Problem, dass Mikrofone die Schallaufnahme ersetzen und somit die richtungsabhängige Analyse erschwert wird. Die grobe Lokalisation, wie sie einem Normalhörenden auch mit nur einem Ohr möglich ist, ist zusätzlich erschwert. Da es sich bei einer Versorgung mit zwei Hörhilfen um zwei unabhängige Systeme handelt, sind die Signale an beiden Ohren unterschiedlich. Die Hörhilfen arbeiten in der Regel unabhängig voneinander und können die Anforderungen für binaurales Hören nur unzureichend erfüllen (Müller-Deile, 2008). Daher ist die Wiederherstellung „echten“ binauralen Hörens kaum möglich.

Wie bereits beschrieben, zeigen Studien, dass sowohl bimodal als auch bilateral versorgte CI-Träger Verbesserungen im Sprachverstehen im Störschall zeigen und eine Verbesserung des Richtungshörens und der Lokalisationsfähigkeit eintritt (Ching et al., 2007).

Außerdem existiert eine Reihe von Studien in komplexen Hörsituationen, die Effekte wie Kopfschatten, Squelch und Summation bei diesen Patientenkollektiven untersuchten. Es zeigte sich, dass der Kopfschatteneffekt der einzige war, der konstant gemessen werden konnte. Sowohl bilaterale CI-Nutzer (Gantz et al., 2002; Müller et al., 2002) als auch bimodale CI-Nutzer (Ching et al., 2004) profitierten von diesem Effekt. Studien zum Squelch- und Summationseffekt zeigten uneinheitliche Ergebnisse. Während in einigen Studien beide CI-Gruppen sowohl vom Squelch- (Müller et al., 2002; Laszig et al., 2004b; Schleich et al., 2004) als auch vom Summationseffekt profitierten (Müller et al., 2002; Laszig et al., 2004b; Schleich et al., 2004; Ching et al., 2007), wiesen andere Studien nur auf kleine und teilweise nicht-signifikante Effekte hin (Gantz et al., 2002; van Hoesel und Tyler, 2003). Die Ergebnisse waren inter-individuell sehr unterschiedlich.

Patienten mit Resthörigkeit im Tieftonbereich (EAS oder bimodal versorgte Patienten) zeigten einen Zugewinn bei der Identifizierung von Sprechern, was ihnen auch bei der Trennung konkurrierender Sprecher half (Turner et al., 2004). Durch die Wahrnehmung der tiefen Frequenzen, übertragen durch das HG, ist es ebenfalls möglich, prosodische Merkmale der Sprache besser wahrzunehmen (Cullington und Zeng, 2011). Die Verbesserung zeigte sich nicht nur für die Sprachmelodie, sondern setzte sich beim Musikhören weiter fort. Verglichen mit der unilateralen CI-Nutzung wurde z.B. die Wahrnehmung der Tonhöhe verbessert (Büchler et al., 2009).

Bezüglich der Prosodie- und Musikwahrnehmung gibt es noch wenige Untersuchungen, die sich speziell auf das Probandenkollektiv der bilateral versorgten CI-Träger konzentrieren. Es wird jedoch angenommen, dass durch die beiden einzelnen CIs auch die Wahrnehmung für prosodische und klangliche Merkmale erhöht wird. Dies soll aus der Tatsache resultieren, dass unterschiedliche spektrale Merkmale durch die beiden CIs übertragen werden. Im Vergleich zeigte sich jedoch, dass bimodal versorgte Patienten in Untersuchungen bzgl. dieser Merkmale (Tonhöhe, Prosodie etc.) bessere Ergebnisse erzielten als bilaterale Patienten (Cullington und Zeng, 2011).

Mit der Frage nach der Bedeutung zeitlicher und spektraler Merkmale für die Vokalerkennung und für die Klangfarbenwahrnehmung bei bimodal und bilateral versorgten CI-Probanden beschäftigten sich Kong et al. (2012). Bezüglich des Zugewinns durch die zweite Hörhilfe gab es einen Zusammenhang zwischen der Vokalerkennung und der Klangfarbenwahrnehmung. Diejenigen Probanden, die durch die zweite Hörhilfe eine verbesserte Vokalerkennung erreichten, wiesen auch eine bessere Klangwahrnehmung auf. Dieser Zugewinn trat bei den bimodal versorgten CI-Nutzern konsistenter auf als bei den bilateral versorgten Nutzern.

## Kapitel 4

### Sprachverstehen im Störgeräusch

In der klinischen Routine werden Sprachverständlichkeitstests durchgeführt, um Art und Grad einer Hörstörung diagnostizieren zu können und die Verordnung und Anpassung von Hörhilfen abzuklären. Um reale Hörsituationen nachzuempfinden, werden nicht nur Sprachtests in Ruhe, sondern insbesondere Sprachtests im Störgeräusch durchgeführt.

Für die klinischen Untersuchungen wird häufig stationäres Rauschen verwendet, durch das z.B. das Sprachverstehen bei Verkehrslärm simuliert und abgeschätzt werden kann. Um das Sprachverstehen in Situationen mit vielen Personen zu rekonstruieren (sog. Cocktail-Party-Situation), werden sprachähnliche Rauschsignale verwendet, die dem Langzeitspektrum von Sprache entsprechen und zusätzlich, durch Amplitudenmodulation, die zeitlichen Merkmale von Sprache berücksichtigen. Tests mit Sprache als Störgeräusch sind bis dato noch nicht fest in der klinischen Routine verankert. Studien zeigen jedoch, dass die Auswirkungen auf das Sprachverstehen bei NH unterschiedlich sind, je nachdem ob Sprachmaskierer oder Rauschmaskierer als Störsignal Verwendung finden (Freyman et al., 1999; Li et al., 2004).

Insgesamt zeigt sich bei Normalhörenden, dass im Vergleich von verschiedenartigen Rauschsignalen das stationäre Rauschen der effektivste Maskierer ist. Dieses Rauschen hat die höchste Energiedichte, so dass die Hörer große SNRs benötigen, um das Sprachsignal zu verstehen. Bei modulierten Maskierersignalen besteht für Normalhörende die Möglichkeit, weniger stark maskierte Bereiche zu nutzen, um das Zielsignal zu verstehen (Fähigkeit des Lückenhörens vgl. Kapitel 2.2.2).

Ein moduliertes, sprachsimulierendes Rauschen maskiert weniger als ein stationäres Rauschen, was weiterhin durch die im Mittel geringere Energiedichte erklärt werden kann (Wagener und Brand, 2005). Im Gegensatz zum modulierten sprachsimulierenden Rauschen ruft ein konkurrierendes Sprachsignal mehr Maskierung hervor. Der Unterschied zwischen den beiden modulierten Maskierern liegt darin, dass das Sprachsignal noch zusätzlich Feinstruktur und die Information des Gesagten beinhaltet. Die zusätzliche höhere

Maskierung durch ein Sprachsignal wird zentralen auditorischen Phänomenen zugesprochen. Bei der Verwendung von Rauschsignalen wird somit eher Maskierung auf der peripheren Ebene erzeugt, während bei der Verwendung von Sprachsignalen zusätzlich die zentrale auditorische Ebene einbezogen wird.

In komplexen Hörsituationen mit Sprache als Störsignal ist eine Trennung von peripherer und zentraler Maskierung nicht möglich, da hierbei immer beide Maskiererarten präsent sind.

### **4.1 Periphere und zentrale Maskierung**

Das Phänomen der Maskierung ist ein grundlegendes Konzept des Hörens, das auf die Mechanismen der Frequenzanalyse des auditorischen Systems zurückgeführt wird.

Masking is thus the opposite of analysis; it represents the inability of the auditory mechanisms to separate the tonal stimulation into components and to discriminate between the presence and the absence of one of them. The degree to which one component of a sound is masked by the remainder of the sound is usually determined by measuring two thresholds (Licklider, 1951: 1005, zitiert aus Kidd, 2008: 146).

Bei der Verarbeitung komplexer Schalle zerlegt das Gehör das Eingangssignal in seine Frequenzanteile in der Größe der sog. Frequenzgruppen. Frequenzgruppen können als Filter angesehen werden, die eine bestimmte Breite (Frequenzgruppenbreite oder kritische Bandbreite) haben. Die kritische Bandbreite bestimmt das Frequenzauflösungsvermögen. Eine Frequenzgruppe entspricht einer Länge von 1,3 mm auf der Basilarmembran. Bei tiefen Frequenzen sind die Frequenzgruppen schmaler als in höheren Frequenzbereichen. Die kritische Bandbreite ist frequenzabhängig. Zwei gleichzeitig dargebotene Töne unterschiedlicher Frequenz werden nur dann als getrennt wahrgenommen, wenn sie in unterschiedliche Frequenzgruppen fallen. Komplexe Signale, die in eine Frequenzgruppe fallen, werden nicht als unterschiedliche Signale wahrgenommen (Böhme und Welzl-Müller, 2005).

Die anfänglichen Studien zu Maskierungsphänomenen haben ihren Ursprung in Experimenten, die von den Bell Laboratorien durchgeführt wurden. Bereits in deren früheren Studien (Wegel und Lane, 1924; Fletcher, 1929) gab es Erkenntnisse, die darauf hinwiesen, dass es zwei unterschiedliche Arten von Maskierung gibt - eine auf der peripheren und eine auf der zentralen Ebene. Periphere Maskierung kann nur dann entstehen, wenn zwei Töne denselben Bereich auf der Basilarmembran erregen (Kidd, 2008). Die Empfindungsschwelle eines Signals bei gleichzeitiger Anwesenheit eines weiteren Signals wird erhöht. Zentrale Maskierung ergibt sich durch den Konflikt von Sensationen im Gehirn und ist wahrscheinlich zu einem gewissen Grad immer präsent.

## Kapitel 4. Sprachverstehen im Störgeräusch

---

Der Begriff der peripheren Maskierung bezieht sich auf Maskierungsphänomene, wie sie durch Modelle der kritischen Bandbreite bzw. der Filter-Bank-Theorie beschrieben werden können. In ihrer einfachsten Form besagt die Filter-Bank-Theorie, dass anhand der Energie, die in die gleichen kritischen Bänder fällt, Maskierung vorhergesagt werden kann. In den anfänglichen Experimenten wurde die Detektion von Tönen bei gleichzeitiger Verdeckung durch Rauschen oder andere Töne untersucht. Dies ruft überlappende Stimulationsmuster in der Cochlea hervor, die als physiologische Basis für Maskierung galten. Als Grundlage zählt ein bekanntes Experiment von Fletcher (1940). Ziel war es, die Detektionsschwelle eines 2 kHz Testtons, der in einem konstanten Rauschen präsentiert wurde, zu messen. Die Detektionsschwelle wurde in Abhängigkeit der Bandbreite des Rauschens gemessen, die frei wählbar war. Dafür wurde der Zielton im Pegel variiert. Die Schwelle stieg mit zunehmender Bandbreite des Rauschens an, erhöhte sich ab einem bestimmten Punkt jedoch nicht mehr. Eine Erklärung dafür war, dass die Basilarmembran eine Bank von Bandpassfiltern sei, jeder mit eigener charakteristischer Frequenz und bestimmter Bandbreite. Fällt die Energie des Zieltons und des Maskierers in den gleichen Filter (bzw. Frequenzgruppe), ergibt sich die Detektionsschwelle durch die Leistung des Rauschens innerhalb dieses Filters (Hansen, 2007).

In späteren Studien wurde angemerkt, dass auch andere Faktoren, die nicht mit überlappenden Stimulationsmustern auf der Basilarmembran in Beziehung standen, das Ausmaß der Maskierung beeinflussen. Die Phänomene auf der peripheren Ebene reichten nicht als Erklärung für Maskierung aus. Carhart et al. (1969) führten eine Reihe von Experimenten zur Sprachverständlichkeit durch. Sie zeigten, dass Hörer durch die Kombination eines Sprachmaskierers mit einem modulierten Rauschen oder durch die Kombination von zwei Sprachmaskierern in höherem Maße gestört wurden, als es durch die einfache Summation der einzelnen Maskierer zu erwarten gewesen wäre. Sie folgerten daraus, dass das Abstrahieren einer Information aus konkurrierenden Signalen, die sinnvolle Sprache enthielten, schwieriger war, als durch die Analyse des Spektrums beider Maskierer berechnet wurde. Dieser Effekt wurde mit dem Begriff perzeptionelle Maskierung beschrieben. Die Annahme war, dass diese Art von Maskierung nur mit Sprache auftreten würde und dies mit semantischen oder linguistischen Faktoren in Sprache zusammenhängen könnte. Verschiedene Experimente mit komplexen Tonsequenzen zeigten später, dass die perzeptionelle Maskierung nicht nur in Verbindung mit Sprache auftrat (Watson et al., 1976). Bei diesen Experimenten ging die Begriffsbildung der peripheren Maskierung zum „Energetic Masking, EM“ und der perzeptionellen / zentralen Maskierung zum „Informational Masking, IM“ über.

### 4.1.1 Informational Masking - Konzept

Die Begriffe „Informational Masking, IM“ und „Energetic Masking, EM“ wurden das erste Mal von Pollack 1975 in seinem Abstract zu einer Präsentation für die Acoustical Society of America verwendet und geprägt. Mit dem bereits vor Pollack von Carhart et al. (1968, 1969) geprägten Begriff der perzeptionellen Maskierung wurden ebenfalls Phänomene bezeichnet, die heute unter den Begriff IM fallen.<sup>3</sup>

Eine klare Definition für das Phänomen IM zu finden, ist schwierig. Forscher interpretierten IM als diejenige Art von Maskierung, die über die rein energetische Maskierung hinausgeht. Die Schwierigkeit in der Definition besteht immer darin, dass die Grundlage das Phänomen EM bildet (Kidd et al., 2008).

Eine für diese Arbeit gültige Definition versucht, zwischen EM und IM zu unterscheiden, und kann sowohl für sprachliche als auch nicht-sprachliche Maskierer verwendet werden. Nach Leek et al. (1991) beschreibt IM die Störwirkung die Kontextkomponenten auf die Detektion und Diskrimination einer Zielkomponente in einem komplexen Schall haben. Die Störwirkung wird hierbei nicht bzw. nicht allein durch deren spektrale Energie verursacht (EM), sondern durch die zusätzliche Information der Kontextkomponenten. Diese Information ist für den Hörer jedoch nicht von Bedeutung und hat somit eine störende Wirkung auf das Sprachverstehen.

IM beschreibt Phänomene, die oberhalb des peripheren Hörsystems stattfinden und mit zentralen Aspekten des Hörens erklärt werden. Prozesse der Aufmerksamkeit, der auditiven Segregationsfähigkeit oder Gedächtnisleistungen sind Faktoren, die sowohl eng mit der Entstehung als auch mit der Reduzierung von IM verbunden sind (Kidd, 2008). Die Möglichkeit auditorische Gruppen zu bilden, um somit die Information eines Zielsignals zu erkennen und diesem folgen zu können, gehört ebenfalls zu den Aspekten, die beim IM berücksichtigt werden müssen (Bregman, 1990).

IM ist primär ein zentral auditorisches Phänomen, was durch Studien mit kontralateraler Stimulation aufgezeigt werden konnte. Brungart und Simpson (2002) arbeiteten mit Sprachsignalen und präsentierten ihren Hörern eine Überlagerung von Ziel- und Maskierersprecher auf dem einen Ohr (Ziel) und ein weiteres unabhängiges Maskierersignal auf dem kontralateralen Ohr (Maskierer). Handelte es sich bei dem kontralateralen Maskierersignal um Rauschen, waren die Hörer in der Lage, die überlagerten Sprecher, die dem Zielohr präsentiert wurden, zu trennen, und verstanden den Zielsprecher ebenso gut wie ohne Maskierungsgeräusch auf der kontralateralen Seite. Wurde auf dem kontralateralen

---

<sup>3</sup> Einen guten Überblick über das Thema „Informational Masking“ gibt: Kidd, Jr. G., Mason, C.R., Richards, V.M., Gallun, F.J. und Durlach, N.I. (2008). Informational Masking. In Yost, W. A., A.N. Popper und R.R Fay (Hrsg.): Auditory Perception of Sound Sources, New York, Springer, 143 - 189.

Ohr jedoch Sprache dargeboten, war die Trennung von Ziel- und Maskierersprecher im Zielohr schlechter. Die Verschlechterung des Sprachverstehens wurde dem IM zugerechnet und könnte allgemeine Schwierigkeiten bei Verarbeitungsprozessen mit IM widerspiegeln. Die Erhöhung von IM konnte auch in Studien mit nicht-sprachlichen Maskierern beobachtet werden (Kidd et al., 2003).

Völlig losgelöst von der Frage nach der peripheren (Vor-)verarbeitung kann IM nicht betrachtet werden. Die Bedeutung der peripheren Vorverarbeitung wird anhand von Studien klar, in denen die Auswirkungen von IM bei NH im Vergleich zu Personen mit sensorineuralen Hörstörungen untersucht wurden. Allgemein führten bei hörgeschädigten Personen die breiteren Frequenzgruppen innerhalb der Cochlea zu einem höheren Anteil von EM. Untersuchungen zeigten, dass das höhere Ausmaß von EM zu einer gleichzeitigen Reduzierung von IM führte. Der verbleibende Anteil von IM konnte im Vergleich mit NH jedoch schlechter verringert werden (Kidd et al., 2003).

### **4.1.2 „Stimulus uncertainty“ und „Target masker similarity“**

Zwei wichtige Effekte, auf denen IM basiert, sind die „stimulus uncertainty“ und die „target-masker similarity“ (Watson, 2005; Brungart, 2001). „Stimulus uncertainty“ beschreibt die Ungewissheit des Hörers, eine bestimmte Komponente dem Ziel- bzw. dem Maskierersignal zuzuordnen. Der Begriff „Target-masker similarity“ bezieht sich auf die Ähnlichkeit zwischen Ziel- und Maskierersignal hinsichtlich der Intensität, zeitlicher Aspekte und spektraler Eigenschaften. Die ersten Untersuchungen zum Effekt der „stimulus uncertainty“ wurden von Watson et al. (1976) durchgeführt. Er untersuchte die Fähigkeit, Änderungen bestimmter Parameter wie Frequenz oder Intensität eines Zieltons in einer Tonsequenz mit unterschiedlichen Maskierertönen in Abhängigkeit von der „stimulus uncertainty“ wahrzunehmen. Durch Manipulation der Maskierertöne wurde der Grad der „stimulus uncertainty“ verändert. Hierbei zeigten sich große Unterschiede in der Fähigkeit der Hörer, die Änderungen dieser Parameter wahrzunehmen. Je größer also die „stimulus uncertainty“ war, desto größer waren die Auswirkungen von IM.

Neff and Green (1987) führten die sog. „Multitone Masking Experimente“ ein, die in vielen späteren Studien genutzt wurden, um IM zu untersuchen.

Neff and Green (1987) generated masker composed of  $N$  randomly selected tonal components, where  $N$  was varied over a wide range. The task was to detect a pure-tone target of fixed and known frequency located within a distribution of potential masker components, for masker samples drawn at random on every presentation. These multitone maskers were essentially sparse samples of components of Gaussian noise, with the density of components the parameter of interest. They found large amounts of masking for small number of components that, according to classical masking theory, should produce very little energetic masking (Kidd et al., 2008: 149).

Der Effekt der „stimulus uncertainty“ kann jedoch nicht unabhängig von dem Effekt der „target-masker similarity“ betrachtet werden. Beide Effekte sind eng miteinander verbunden. In verschiedenen Studien zeigte sich, dass die Detektionsschwelle eines Zieltons in Anwesenheit von gleichzeitig präsentierten Maskierertönen niedriger war, je unähnlicher Ziel- und Maskierertöne waren. Durch eine geringere „target-masker similarity“ nahmen Verwechslungen und Gruppierungen von Ziel- und Maskierersignalen ab (Kidd et. al., 1994; Durlach et. al., 2003). Dies führte ebenfalls zu einer Reduzierung von IM.

Ein dritter Faktor, der explizit sprachlichen Maskierern vorbehalten ist, ist das „linguistic masking“. Dies besagt, dass IM auch dann möglich ist, wenn das Maskiersignal unverständlich ist und keinen semantischen Kontext besitzt, aber sprachliche Merkmale beinhaltet. Das wäre z.B. bei rückwärts präsentierter Sprache oder einer, dem Hörer unbekannt, Fremdsprache der Fall. Solche Maskiersignale enthalten phonetische Merkmale, die mit denen des Zielsignals verwechselt werden können, und somit IM produzieren (Hawley et al., 2004).

### **4.1.3 Informationale Masking mit Sprachsignalen**

Das Phänomen IM gibt es in jeder alltäglichen Kommunikationssituation, in der ein Hörer mehr als einer Person zuhört. Von Cherry (1953) wurde ein solches Szenario mit dem Begriff „cocktail party problem“ bezeichnet. Theoretisch erhält man in einer solchen Kommunikationssituation IM in maximaler Ausprägung, wenn sowohl „target-masker similarity“ als auch „stimulus uncertainty“ besonders hoch sind. In Bezug auf konkurrierende Sprecher würde sich diese Kondition bei stimmlich identischem Ziel- und Maskiersprecher von gleichem Pegel ergeben. Es würden keine stimmlichen Variablen wie eine unterschiedliche  $F_0$  oder unterschiedliche  $F_n$  auftreten, anhand derer die Sprecher voneinander getrennt werden könnten. Hörer wären zwar in der Lage, die Wörter zu verstehen, könnten sie jedoch nicht sicher dem Ziel- oder Maskiersprecher zuordnen. Es würden 50% der Ziel- und 50% der Maskiererwörter genannt werden, d.h. Fehler würden sich dadurch ergeben, dass die Wörter der beiden Sprecher verwechselt werden (Pyschny et al., 2009).

Sobald es einen Unterschied zwischen den Sprechern gibt, gelingt eine bessere Trennung der Sprecher, da eine herabgesetzte „target-masker similarity“ wiederum zu einer reduzierten „stimulus uncertainty“ führt (Brungart et al., 2002).

Um „target-masker similarity“ und „stimulus uncertainty“ zu reduzieren, ist die Einführung eines Pegelunterschiedes (SNR) ein wichtiger Faktor. Handelt es sich um Sprachmaterial, spricht man vom sogenannten TMR für englisch „target-to-masker ratio“. Bezogen auf alltägliche Kommunikationssituationen mit zwei Sprechern ändert sich der TMR, wenn sich



## **Kapitel 4. Sprachverstehen im Störgeräusch**

---

die Sprecher z.B. voneinander weg bewegen. In diesem Fall ist der Hörer in der Lage, aufgrund der unterschiedlichen wahrgenommenen Lautstärke, einen Sprecher zu extrahieren. Normalhörende Personen können sowohl den laueren als auch den leiseren Sprecher verstehen (Brungart, 2001).

Ein weiterer Faktor, der das Sprachverstehen bei konkurrierenden Sprechern beeinflusst, ist die Anzahl der Störsprecher. Die Sprachverständlichkeit sinkt mit steigender Anzahl der konkurrierenden Sprecher. Brungart et al. (2002) untersuchten die Sprachverständlichkeit zwischen keinem und drei Maskierersprechern gleichen Geschlechts bei gleichem Pegel. Das Sprachmaterial wurde einkanalig aufgenommen und NH diotisch über Kopfhörer in Mono-Qualität dargeboten. Mit jedem hinzugekommenen Störsprecher sank die Sprachverständlichkeit um den Faktor 0.6.

Nicht nur die Anzahl der Störsprecher, sondern auch die stimmliche Charakteristik der Sprecher tragen dazu bei, wie gut konkurrierende Sprecher voneinander getrennt werden können. Sprecherstimmen können anhand ihrer  $F_0$ , ihrer  $F_n$ , ihrer Sprechgeschwindigkeit, ihrer Betonung, ihrem Akzent etc. voneinander getrennt werden (Brungart et al., 2002). Studien haben aufgezeigt, dass das Sprachverstehen bei mehreren Sprechern am besten ist, wenn diese unterschiedlichen Geschlechts sind. Die Sprecher können meist anhand der unterschiedlichen  $F_0$  voneinander unterschieden und getrennt werden (Brungart et al., 2001). Ein Unterschied in  $F_0$ , der  $1/6$  einer Oktave entspricht, ist ausreichend, um eine signifikante Verbesserung im Sprachverstehen zu erreichen (Assmann und Summerfield, 1990).

Ein weiterer Faktor, der zu einer besseren Trennung von Ziel- und Maskierersprecher beiträgt, ist die räumliche Trennung. Bei den bisher erwähnten Untersuchungen wurden verschiedene Sprechersignale alle aus derselben Richtung dargeboten. Dies ist in normalen Hörsituationen nicht der Fall. Sprecher sind räumlich voneinander getrennt, so dass unterschiedliche Signale an beiden Ohren ankommen (vgl. Kapitel 2). Wenn ein konkurrierender Sprecher räumlich vom Zielsprecher getrennt ist, kann sich das Sprachverstehen um nahezu 25% erhöhen, im Vergleich zu einer Situation in der Ziel- und Maskierersprecher räumlich nicht getrennt sind. Bei mehreren konkurrierenden Sprechern können sich prozentual die korrekten Antworten durch räumliche Trennung sogar verdoppeln (Brungart und Simpson, 2002).

Ein ebenfalls wichtiger Punkt ist die Vorinformation, die die Hörer darüber haben, wo sich der Zielsprecher befindet und wer der Zielsprecher ist (Kidd et al., 2005).

### 4.1.4 Exemplarische Untersuchungen mit Normalhörenden

Im englischsprachigen Raum hat sich der „coordinate response measure, CRM“-Test (Bolia et al., 2000) als effektives Testwerkzeug bewährt, um Effekte des IM zu erforschen. Hiermit werden vornehmlich Effekte untersucht, die entstehen, wenn Sprache als Störgeräusch verwendet wird. Ursprünglich wurde der CRM-Test für den militärischen Bereich konzipiert, um die Schwierigkeiten aufzuzeigen, mit denen Jet-Piloten konfrontiert werden, wenn unterschiedliche, sprachliche Information gleichzeitig dargeboten wird (Moore, 1981).

Dieser Test wird nun in einer Vielzahl von Studien eingesetzt, um die Sprachverständlichkeit in Situationen mit konkurrierenden Sprechern zu untersuchen (z.B. Brungart, 2001). Der CRM-Test wird als closed set, forced-choice Methode durchgeführt. Die Satzstruktur der dargebotenen Sätze des CRM-Tests ist immer „Ready [call sign] go to [color] [number] now“. Das Schlüsselwort (call sign) des Zielsatzes ist der Name „Baron“. Die Aufgabe der Hörer ist, dem Satz zu folgen, der mit dem Namen „Baron“ beginnt und die dazugehörige Farbe und Zahl zu nennen. Gleichzeitig wird dieser Zielsprecher durch einen weiteren Maskierersprecher überlagert, dessen Sätze gleich strukturiert sind, jedoch andere Testwörter enthalten.

Brungart (2001) untersuchte mit Hilfe des CRM-Tests die Effekte von IM und EM bei normalhörenden Personen. Um zwischen IM und EM zu differenzieren, verwendete er unterschiedliche konkurrierende Sprecher und moduliertes, sprachsimulierendes Rauschen als Maskierersignale. Die Sprachverständlichkeit sank bei Verwendung des modulierten Rauschens aufgrund ungünstiger TMRs. Bei Verwendung eines Sprachmaskierers entstand die größere Schwierigkeit bei der Differenzierung stimmlicher Charakteristiken im Vergleich zu ungünstigen TMRs. Der Grund hierfür war der beim IM wesentliche Effekt der „target-masker similarity“. Brungart (2001) untersuchte in dieser Studie erstmals die Fehlerverteilung. Da durch die spektrale Überlagerung (EM) Wörter unhörbar werden können, zog er den Schluss, dass falsch oder nicht verstandene Wörter durch EM hervorgerufen würden. Hingegen sind Fehler, die sich aufgrund von Verwechslungen zwischen Ziel- und Maskierersprecher ergeben (z.B. „red“ statt „blue“) und die durch die Charakteristiken „stimulus uncertainty“ und „target-masker similarity“ hervorgerufen werden, dem IM zuzurechnen. Hierbei entsteht ein Problem sowohl bei der Bildung auditorischer Gruppen als auch bei der Fähigkeit, einzelne auditorische Ströme zu formen (Shinn-Cunningham und Best, 2008).

For sequences of related sounds, or “streams,” it can also occur that the connection between successive elements is lost because of another (masking) stimulus. Connected speech is one such example, in which the ongoing comprehension of information requires maintaining the integrity of the perceptual stream of sounds emanating from a specific sound source (Kidd et al., 2008: 182).

## **Kapitel 4. Sprachverstehen im Störgeräusch**

Verwechslungsfehler werden als ein Phänomen gestörter „Stream segregation“ interpretiert (s. Kapitel 2). Der Zielsatz kann nicht als eigener auditorischer Strom von dem Maskierersatz separiert werden. Gründe dafür können z.B. sein, dass der stimmlichen Charakteristik des Zielsprechers nicht gefolgt werden kann oder dass die Aufmerksamkeit in eine falsche Richtung (im Fall von räumlich getrennten Sprechern) gelenkt wird (Kidd, 2008).

In einer weiteren Studie untersuchten Brungart et al. (2001) die Fähigkeit Normalhörender, einen Sprecher zu verstehen, wenn mehr als ein Maskierersprecher präsent war. Im Gegensatz zu der vorherigen Studie mit nur einem Sprachmaskierer zeigte sich, dass ungünstige TMRs das Sprachverstehen verschlechterten. Hinsichtlich der Stimmunterschiede war das Sprachverstehen der Zielsätze am besten, wenn die Maskierersätze von einem Sprecher anderen Geschlechts gesprochen wurden. Hingegen konnte der Zielsatz nur sehr schlecht verstanden werden, wenn Ziel- und Maskierersprecher gleich waren. Bei negativen TMRs war das Sprachverstehen abhängig von der Anzahl der Maskierersprecher. Das Sprachverstehen war bei zwei oder drei konkurrierenden Sprechern schlechter als bei einem einzelnen konkurrierenden Sprecher.

Eine Vielzahl von Studien hat sich mit der Trennung konkurrierender Sprecher beschäftigt und hierbei besonders den Unterschied zwischen männlichen und weiblichen Stimmen untersucht. Brox und Nootboom (1982) verwendeten als erste denselben Sprecher als Ziel- und Maskierer. Sie manipulierten die Maskiererstimme, indem sie F<sub>0</sub>-Unterschiede einfügten. Das Verstehen des Zielsprechers verbesserte sich mit zunehmender Differenz in F<sub>0</sub>.

Für die Erkennung und Diskrimination von Sprechern unterschiedlichen Geschlechts spielt nicht nur F<sub>0</sub> eine Rolle. Darwin et al. (2003) haben zusätzlich die Vokaltraktlänge (VTL) verändert, was Änderungen in den Formantfrequenzen (F<sub>n</sub>) hervorruft. Sie haben diese beiden Parameter im Hinblick auf die Trennung konkurrierender Sprecher untersucht. Unterschiede in F<sub>0</sub> verbesserten das Sprachverstehen, wohingegen Unterschiede in VTL nur wenig Verbesserung erzielten. Die kombinierte Änderung von F<sub>0</sub> und VTL führte zu einem super-additiven Effekt, d.h. die Ergebnisse, die durch die separaten Änderungen von F<sub>0</sub> und VTL erzielt wurden, wurden übertroffen.

In Pyschny et al. (2009) wurde als Ausgangskonfiguration ebenfalls derselbe Ziel- und Maskierersprecher verwendet. Die Stimme des Maskierersprechers wurde künstlich in F<sub>0</sub> verändert. Weiterhin wurden Pegelunterschiede (TMR) zwischen Ziel- und Maskierersprecher eingefügt. Es konnte gezeigt werden, dass bei NH sowohl die kleinsten getesteten Unterschiede von 20 Hz in der Grundfrequenz als auch die kleinsten getesteten Pegelunterschiede von 5 dB zu einem verbesserten Sprachverstehen führten. Anlehnend an die Studie von Brungart (2001) wurde eine Analyse der Fehler vorgenommen. Der Großteil der Fehler konnte auf Verwechslungen von Ziel- und Maskiererwörtern zurückgeführt

## Kapitel 4. Sprachverstehen im Störgeräusch

---

werden. Bei gleichem Ziel- und Maskierersprecher wurden 49% der Wörter des Zielsatzes und 49% der Wörter des Maskierersatzes verstanden. Mit Unterschieden in F0 oder TMR verringerten sich die Verwechslungen. Nur wenige Fehler resultierten daraus, dass Wörter falsch oder gar nicht verstanden wurden. Das spricht dafür, dass die eingefügten Unterschiede zwischen Ziel- und Maskierersprecher bei Normalhörenden vornehmlich zu einer Reduktion von IM führten.

Hinsichtlich der räumlichen Trennung von Ziel- und Maskierersignal zeigen Studien, dass die Sprachverständlichkeit durch den Kopfschatten, binaurale Prozesse und bessere Separation der auditorischen Ströme ansteigt.

Arbogast et al. (2002) untersuchten den Effekt der räumlichen Quellentrennung mit unterschiedlichen Maskierertypen, die den Bereich von EM bis IM abdeckten. Ziel- und Maskierersätze wurden mithilfe eines CI-Simulationsprogramms modifiziert. Das Sprachmaterial des CRM-Tests wurde in 15 Frequenzbänder gefiltert. Als Maskierer wurde entweder das Sprachmaterial des CRM-Tests oder Schmalbandrauschen verwendet, welche ebenfalls in 15 Frequenzbänder gefiltert wurden. Danach erfolgte die Erstellung von drei unterschiedlichen Ziel-Maskierer-Konfigurationen.

(1) Different-band sentence (DBS) masker: Acht Frequenzbänder des Zielsatzes wurden mit sechs Frequenzbändern des Maskierersatzes kombiniert. Hierbei waren die Bereiche der Frequenzbänder unterschiedlich, so dass keinerlei Überlappung im Frequenzbereich stattfand und vornehmlich IM produziert werden sollte.

(2) Different-band noise (DBN) masker: Anstatt der sechs Frequenzbänder des Maskierersatzes wurden sechs Frequenzbänder mit sprachsimulierendem Rauschen eingefügt. Im Vergleich zur Konfiguration (1) sollte hier der Anteil des IM gering sein und der Anteil des EM, aufgrund der fehlenden Überlappung, unverändert. Konfiguration (2) sollte als Kontrolle für den Anteil von EM in Konfiguration (1) dienen.

(3) Same-band noise (SBN) masker: Acht Frequenzbänder des Zielsignals wurden mit den gleichen acht Frequenzbändern eines Rauschsignals überlagert. Hierbei entstand eine Konfiguration, durch die ausschließlich EM produziert werden sollte.

Das Zielsignal wurde aus einem Lautsprecher direkt vor dem Hörer (0° Azimut-Winkel) dargeboten, während das Maskierersignal entweder ebenfalls von vorne oder um 90° nach rechts verschoben, dargeboten wurde. Die Ergebnisse zeigten Verbesserungen durch die Quellentrennung, die für den „Informational Masker“ (DBS) größer waren als für die „Energetic Masker“ (DBN und SBN). In einer weiterführenden Studie (Arbogast et al., 2005) wurde die Auswirkung der Quellentrennung mit demselben Messaufbau zusätzlich mit Hörgeschädigten untersucht. Auch wenn sie weniger als die NH von der Quellentrennung profitierten, war die Verbesserung im Sprachverstehen für einen „Informational Masker“ größer als für einen „Energetic Masker“.

## **Kapitel 4. Sprachverstehen im Störgeräusch**

---

Nicht explizit im Kontext IM untersuchten Hawley et al. (2004) das Sprachverstehen in einem Cocktail-Party Szenario unter Berücksichtigung binauralen Hörens. Die Sprachverständlichkeitsschwellen von Sätzen wurden in Anwesenheit von 1 - 3 Störquellen gemessen. Entweder handelte es sich um denselben Sprecher, der auch den Zielsatz sprach, um rückwärts dargebotene Sprache desselben Sprechers, stationäres sprachsimulierendes Rauschen oder sprachsimulierendes Rauschen, das durch die Einhüllende der Sätze moduliert wurde. Die Maskierer wurden aus unterschiedlichen Richtungen dargeboten. Die Ergebnisse wiesen auf einen höheren Zugewinn durch binaurales Hören hin, wenn es sich um stimmhafte, mit einer Grundfrequenz versehene Maskierer (gleicher Sprecher und rückwärts dargebotene Sprache) handelte, unabhängig von der Richtung, aus der das Maskierersignal präsentiert wurde. Bei Nutzung nur eines Maskierers verbesserte sich die Sprachverständlichkeitsschwelle, bei mehreren Maskierern hingegen änderten sich die Sprachverständlichkeitsschwellen nicht mehr. Das lag daran, dass bei mehreren Störsignalen kein Lückenhören mehr möglich war. Je mehr Maskierersignale vorkommen, desto eher werden die Lücken, die für das Verstehen des Zielsatzes genutzt werden könnten, durch die Energie weiterer Maskierersignale gefüllt.

Den wichtigen Aspekt der Vorabinformation im Kontext IM untersuchte Freyman et al. (2004). Unter anderem wurde der Zielsprecher präsentiert, bevor ein verdeckendes Störgeräusch hinzukam. Das Verstehen des letzten Wortes wurde überprüft, das nicht in der Vorabpräsentation enthalten war, später aber mit einem Maskierer überdeckt wurde. Verbesserungen zeigten sich nur in der Kondition, in der Ziel- und Maskierersprecher aus der gleichen Richtung (vorne, 0° Azimut-Winkel) präsentiert wurde. Diese Verbesserung blieb konstant, auch wenn der Satz vorher von einem anderen Sprecher gesprochen wurde bzw. den Hörern schriftlich zu lesen angeboten wurde. In einer aktuelleren Studie (Jones und Freyman, 2012) wurde nach „Energetic Maskern“ und „Informational Maskern“ aufgeteilt. Die Probanden konnten die Vorabinformation bei einem „Energetic Masker“ gut nutzen, um Sprache besser zu verstehen. Die Ergebnisse mit einem „Informational Masker“ waren weniger eindeutig. Es wurde jedoch angenommen, dass IM reduziert werden könnte, sobald die Aufmerksamkeit der Hörer auf das Zielsignal gelenkt würde.

### **4.1.5 Exemplarische Untersuchungen mit Cochlea-Implantat-Trägern**

Die beispielhaft vorgestellten Studien zeigen auf, dass Unterschiede zwischen den Sprecherstimmen, Pegelunterschiede und Unterschiede in der räumlichen Aufteilung der Sprecher sowie Vorabinformation über die Sprecher oder das Gesprochene von NH genutzt werden können, um den Zielsprecher besser zu verstehen. Die Folge ist, dass IM reduziert

wird. Die bis dato durchgeführten Studien mit CI-Trägern zeigen, dass die Ergebnisse der NH nicht auf das Probandenkollektiv der CI-Träger übertragen werden können.

Insgesamt muss bedacht werden, dass das Sprachverstehen im Störgeräusch für CI-Träger erschwert ist. Das liegt an der fehlenden Fähigkeit des „Lückenhörens“ und an der eingeschränkten Trennung auditorischer Ströme. Da die periphere Ebene gestört ist, müssen, auf der für das IM wichtigen zentralen Ebene, bereits veränderte Informationen verarbeitet werden.

In vielen bisher durchgeführten Studien wurden Sprachmaskierer verwendet, um das Sprachverstehen von CI-Trägern in Kommunikationssituationen mit mehreren Sprechern zu untersuchen. Die Frage nach den Auswirkungen von IM bei CI-Trägern stand jedoch nur bei wenigen Studien explizit im Fokus. In den meisten Studien wurden entweder immer die gleichen Maskierersätze verwendet (z.B. Kong et al., 2005), was den Effekt der „stimulus uncertainty“ minimierte, oder Sprache wurde rückwärts angeboten (z.B. Kong und Carlyon, 2007), was den Informationsgehalt und somit den Effekt der „target-masker similarity“ verringerte. In solchen Fällen war die irrelevante Information, durch die IM gekennzeichnet ist, eher klein.

Stickney et al. (2004) untersuchten das Sprachverstehen von CI-Trägern bei konkurrierenden Störsprechern. Als Zielsprecher diente eine männliche Stimme. Als Maskierer wurde entweder der gleiche Sprecher, ein anderer männlicher Sprecher oder eine Sprecherin verwendet. Das Sprachverstehen wurde bei unterschiedlichen TMRs gemessen. Im Gegensatz zum Sprachverstehen in Ruhe war das Sprachverstehen in einer Situation mit konkurrierenden Sprechern stark herabgesetzt. Eine Erklärung dafür war, dass mehr IM entstehe, da die limitierte spektrale Auflösung im CI zu einer größeren „target-masker similarity“ führe. Steigende TMRs verbesserten das Sprachverstehen in geringem Maße.

Zusätzlich zu konkurrierenden Sprechern (männlich oder weiblich) setzten Qin und Oxenham (2003) auch stationäres sprachsimulierendes Rauschen und moduliertes sprachsimulierendes Rauschen ein. Die Stimuli wurden mithilfe eines CI-Simulationsprogramms in 4, 8 oder 24 Frequenzbänder gefiltert und Normalhörenden dargeboten. Untersuchungen solcher Art lassen Rückschlüsse auf die Performanz realer CI-Träger zu. Zum Vergleich wurde auch eine ungefilterte Version der Stimuli getestet. Bei der CI-Simulation waren die Sprachverständlichkeitsschwellen bei Verwendung fluktuierender Störsignale schlechter als bei Verwendung des stationären Störsignals. Als Erklärung galt die reduzierte Darbietung von F0, die normalerweise durch tiefe Frequenzen und durch die Feinstruktur übertragen wird.

Übergehend zu Studien mit beidseitig versorgten CI-Trägern oder anderen Versorgungsformen konnte vor allem bei bimodal versorgten CI-Trägern bzw. EAS-versorgten CI-Trägern gezeigt werden, dass eine verbesserte Auflösung und Übertragung

der tiefen Frequenzen durch ein HG zur Identifizierung von Sprechern und damit zur Unterscheidung von konkurrierenden Sprechern beiträgt (Turner et al., 2004).

Kong et al. (2005) untersuchten, inwieweit vorhandenes Tiefton-Restgehör auf der nicht-implantierten Seite zusätzliche Information überträgt, die dem CI-Nutzer hilft, das Sprachverstehen zu verbessern. Sie untersuchten das Verstehen mit dem HG alleine, dem CI alleine und in der Kombination CI + HG in einer Situation mit konkurrierenden Sprechern. Die Probanden wiesen mit dem HG alleine keinerlei Sprachverstehen auf. Trotzdem half das HG in der Kombination mit dem CI. Das Sprachverstehen stieg, im Vergleich mit dem CI alleine, an. Dies galt besonders dann, wenn sich F0 und die spektralen Charakteristiken des Maskierersprechers stark von denen des Zielsprechers unterschieden. Ein bimodaler Zugewinn ergab sich weiterhin durch Pegelunterschiede zwischen den Sprechern.

Es existiert eine Vielzahl von Studien, die sich mit dem Vorteil des binauralen Hörens und der räumlichen Trennung von Ziel- und Störsignal befassen. Allerdings besteht auch hier das Problem, dass es sich nicht explizit um Untersuchungen zum IM handelt.

Garadat et al. (2009) führten eine Studie mit NH durch, in der sie die Stimuli mit einem CI Simulationsprogramm modifizierten. Die Stimuli wurden räumlich getrennt. Der größte Zugewinn zum Sprachverstehen durch die räumliche Trennung wurde erreicht, wenn ein hoher Anteil von IM vorlag. Wurden ohne räumliche Trennung Ziel- und Maskierersignale häufig verwechselt, verringerten sich die Verwechslungen mit der räumlichen Trennung.

In der Annahme, dass die Vorteile des binauralen Hörens auch bei CI-Trägern von der Art des Maskierers abhängig sind, untersuchten Loizou et al. (2009) das Sprachverstehen bilateral versorgter CI-Träger in einem „Cocktail-Party Szenario“. Basierend auf der Studie von Hawley et al. (2004) wurden gleichzeitig mehrere (1 bis 3) Störsignale von unterschiedlichen Positionen präsentiert. Bei Verwendung der Rauschmaskierer war das Sprachverstehen der CI-Träger ähnlich dem der NH. Bei Verwendung von Sprachmaskierern konnten binaurale Prozesse kaum genutzt werden. Insgesamt war das Sprachverstehen der CI-Träger jedoch unabhängig vom Maskierertyp.

## KAPITEL 5

### Zusammenfassung

Wie in Kapitel 2 beschrieben, übernimmt das auditorische System eine der Hauptfunktionen bei dem Verstehen von Sprache und ist für die menschliche Kommunikation essentiell. Erst durch die genaue Analyse eines akustischen Reizes und dessen Verarbeitung und Weiterleitung ist der Mensch in der Lage, sprachliche Signale von nicht-sprachlichen Signalen zu unterscheiden. Wird ein Reiz als Sprachsignal erkannt, findet in weiterführenden Prozessen die Zuweisung der Bedeutung für dieses Signal statt. Das Gehör reagiert auf sprachliche Signale besonders sensibel. So ist es Normalhörenden möglich, bei gleichzeitigem Auftreten sprachlicher und nicht-sprachlicher Signale, das Sprachsignal in einer solchen Situation herauszufiltern. Wahrscheinlich sind es gemeinsam auftretende akustische Gegebenheiten, die dafür sorgen, dass ein Sprachsignal von einem Laut bis hin zu Sätzen zu einem gemeinsamen Objekt zusammengefasst wird. Dadurch ist es möglich, aufeinanderfolgende Sprachlaute zu einem fortlaufenden Strom zu gruppieren. Gleichzeitig auftretende Signale können voneinander separiert werden, so dass dem, für einen Hörer wichtigen, sprachlichen Signal gefolgt werden kann. Schwieriger wird die Trennung von gleichzeitig auftretenden Signalen, wenn es sich bei allen Signalen um Sprache handelt. Allerdings sind auch hier Normalhörende in der Lage, sprecherspezifische Merkmale herauszufiltern, so dass es möglich ist, einem Sprecher in einer Situation mit mehreren Sprechern zu folgen.

Das auditorische System kann jedoch aufgrund unterschiedlichster Ursachen in seiner Funktion gestört werden. Dies hat erhebliche Auswirkungen auf die Kommunikation und kann die Betroffenen auf der sozialen Ebene stark einschränken.

Es können verschiedene Arten von Hörstörungen entstehen. Die zwei häufigsten Arten von Hörstörungen werden unter den Begriffen Schallleitungsschwerhörigkeit und Schallempfindungsschwerhörigkeit zusammengefasst. Das Außen- und das Mittelohr betreffend kann die Schallleitungsschwerhörigkeit entstehen, die jedoch durch apparative



und operative Therapien, meist mit wenigen verbleibenden Einschränkungen, behoben werden kann.

Im Vergleich dazu steht die Schallempfindungsschwerhörigkeit, die ihren Ursprung im Innenohr hat. Hierbei stoßen die Therapien an ihre Grenzen. Den größten Teil der apparativen Therapie macht die Versorgung mit Hörgeräten aus. Je größer jedoch der Hörverlust, desto schwieriger ist es, Sprachverstehen und Klangerlebnis ähnlich wie das eines Normalhörenden „wieder“ zu erlangen. Häufig gelingt das Sprachverstehen in ruhiger Umgebung. Eine Kommunikation im Störgeräusch oder mit mehreren Sprechern ist jedoch kaum möglich. Der Grund hierfür ist, dass durch die gestörte, periphere Analyse des Schallsignals nicht alle wichtigen Komponenten des Sprachsignals fehlerfrei verarbeitet werden können. Somit wird Sprache nicht mehr eindeutig als Sprache erkannt. Die Zusammenfassung relevanter auditorischer Objekte gelingt nicht mehr. Wenn der Hörverlust so stark ist, dass auch mit Hörgeräten kein Sprachverstehen mehr möglich ist, ist nach sorgfältig ausgesuchten Indikationskriterien die Versorgung mit einem Cochlea-Implantat eine Lösung für eben solche Patienten. Das in Kapitel 3 beschriebene Cochlea-Implantat, als Prothese, übernimmt die Funktionen des Innenohres. Es besteht aus zwei Hauptkomponenten, dem Implantat und dem Sprachprozessor. Der Sprachprozessor ist für die Umwandlung der akustischen Signale aus der Umwelt in elektrische Signale zuständig. Diese Signale werden dann induktiv über eine transkutane Verbindung mittels einer Sendespule zum Implantat übertragen. Die Umwandlung akustischer in elektrische Signale erfolgt aufgrund speziell entwickelter Sprachsignalverarbeitungsstrategien. Durch die Verarbeitungsstrategie soll ein möglichst natürliches Reizmuster gebildet werden, das den Hörnervenfasern dargeboten wird. Das Sprachverstehen in Ruhe ist für die Mehrzahl der mit einem Cochlea-Implantat versorgten Patienten durchaus auch ohne Mundbild möglich. Trotz voranschreitender Technik sehen sich jedoch die meisten Cochlea-Implantat-Nutzer in einigen Situationen vor große Probleme gestellt. So ist vor allem das Sprachverstehen mit Störsignalen (Energetic und Informationale Masking) ein Hauptaspekt (s. Kapitel 4). Ähnlich wie bei apparativ unversorgten Hörgeschädigten, als auch bei den mit einem Hörgerät versorgten Patienten, ist die reduzierte periphere Analyse der akustischen Signale eingeschränkt. Das führt zu einer fehlerhaften Zusammenfassung auditorischer Objekte. Zusätzlich werden durch technische Einschränkungen nicht alle akustischen Merkmale hinreichend übertragen. So wird die, für das Verstehen im Störgeräusch, wichtige Feinstruktur in Signalen und die, für die Erkennung einer Stimme, wichtige Grundfrequenz kaum übertragen. Falls Patienten über Restgehör verfügen, können in solchen komplexen Hörsituationen ein Hörgerät auf der kontralateralen Seite des Cochlea-Implantates (bimodale Versorgung) oder eine elektro-akustische Stimulation das Sprachverstehen verbessern.

## **Kapitel 5. Zusammenfassung**

---

Ergibt sich dadurch kein Vorteil im Sprachverstehen, stellt sich eventuell die Frage nach einer bilateralen Versorgung (zwei Cochlea-Implantate).

## Kapitel 6

### Zielsetzung und Hypothesen

“This is indeed a mystery,” I remarked. “What do you imagine that it means?”  
“I have no data yet. It is a capital mistake to theorise before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts.”

(Sir Arthur Conon Doyle, Sherlock Holmes, A scandal in Bohemia)

Anlehnend an Studien mit Normalhörenden (NH) werden im Folgenden unterschiedliche Parameter untersucht, die zur Trennung konkurrierender Sprecher bei Cochlea-Implantat (CI)-Trägern beitragen können. Für die Experimente wird die Stimme des Maskierersprechers aus der des Zielsprechers generiert. Die zwei dominantesten sprecherspezifischen Merkmale, die Grundfrequenz ( $F_0$ ) und die Formantfrequenzen ( $F_n$ ), werden bei der Maskiererstimme künstlich verändert. Zeitliche und prosodische Variationen der Stimme werden beibehalten. Die Änderungen der Parameter sollen auch bei CI-Trägern zu einer Reduzierung von Informationale Masking führen. Es bleibt aufzuzeigen, ob Energetic Masking (EM) und Informationale Masking (IM) ähnliche Auswirkungen auf das Sprachverstehen haben wie bei NH oder ob ähnlich wie bei Hörgeschädigten, die nicht mit einem CI versorgt sind, EM dominiert und IM eine untergeordnetere Rolle spielt. Zusätzlich zu den sprecherspezifischen Merkmalen werden Unterschiede im Pegel zwischen den Sprechern und die unterschiedliche räumliche Aufteilung der Sprecher in die Untersuchungen mit einbezogen.

#### 6.1 Vortest: Sprecherunterscheidung und Sprechererkennung

Als Grundlage für die drei Experimente zum IM bei CI-Trägern dient ein Vortest. Dieser beschäftigt sich mit der Erkennung und Diskrimination von Sprecherstimmen auf Grundlage von Änderungen in  $F_0$  und in  $F_n$ . Ziel ist herauszufinden, ob Unterschiede zwischen Sprechern, die *sequentiell* präsentiert werden, wahrgenommen werden können. Erst wenn das der Fall ist, wird untersucht, ob diese Fähigkeiten in einer Kommunikationssituation mit *simultanen* Sprechern genutzt werden können, um diese voneinander zu trennen. Durch die

Trennung sollte sich das Sprachverstehen in diesen komplexen Kommunikationssituationen verbessern.

Anlehnend an die Studien von Darwin et al. (2003) und Fu et al. (2005) wird Hypothese 1 aufgestellt:

*H1: Die Sprechererkennung und -diskrimination aufgrund von F0-Unterschieden gelingt besser als die aufgrund von Unterschieden in den  $F_n$*

### **6.2 Experiment 1: Pegel- und Grundfrequenzunterschiede**

Das erste Experiment dient dazu festzustellen, inwieweit unilateral versorgte CI-Träger in der Lage sind, einen Zielsatz zu verstehen, wenn dieser durch einen simultan dargebotenen Maskierersatz überlagert wird. Weiterhin soll aufgezeigt werden, welche Art von Unterschied zwischen dem Ziel- und dem Maskierersprecher von CI-Trägern genutzt wird, um ein besseres Verstehen des Zielsprechers zu erlangen. Unterschiedliche Pegel ("target-to-masker ratio", TMR) und künstlich eingefügte Unterschiede in der Grundfrequenz ( $\Delta F_0$ ) sollen einen Hinweis darauf geben, ob diese Unterschiede die Maskierung, IM und EM, beeinflussen. Außerdem dient die gezielte Veränderung dazu, einzelne Parameter unabhängig voneinander untersuchen zu können.

Studien zeigten, dass Pegelunterschiede (TMR) von CI-Trägern genutzt werden können (Stickney et al., 2004). Da explizite Ortskodierung in den meisten CI-Systemen im Bereich von F0 nicht gegeben ist, wird die Information über F0 hauptsächlich durch temporale Merkmale übertragen und daher unvollständig verarbeitet. Aufgrund bisheriger Studien werden folgende Hypothesen aufgestellt:

*H2: Die Sprachverständlichkeit verbessert sich durch das Einfügen von Pegelunterschieden zwischen Ziel- und Maskierersprecher. Sowohl EM als auch IM werden reduziert.*

*H3: Die Wahrnehmung von Unterschieden in der Grundfrequenz zwischen den Sprechern kann nur eingeschränkt zur Unterscheidung konkurrierender Sprecher genutzt werden. EM und IM werden durch F0-Unterschiede nur wenig reduziert.*

### **6.3 Experiment 2: Grundfrequenz- und Formantfrequenzunterschiede**

Da das CI die F0 und spektrale Merkmale wie z.B. Formantfrequenzen ( $F_n$ ) nur eingeschränkt überträgt, beschäftigt sich das zweite Experiment mit dem eventuellen Vorteil

## Kapitel 6. Zielsetzung und Hypothesen

---

durch ein zusätzliches, kontralaterales Hörgerät (HG). Da Resthörigkeit bei den meisten Patienten eher im tieffrequenten Bereich besteht, können diese Frequenzen noch ausreichend durch ein HG verstärkt werden. Die Probanden in Experiment 2 sind bimodal mit CI und HG versorgt.

Darwin et al. (2003) konnten aufzeigen, dass Unterschiede in F<sub>0</sub> zwischen konkurrierenden Sprechern von NH genutzt werden können, so dass der Zielsprecher besser verstanden wird. Änderungen der F<sub>n</sub> des Maskierersprechers verbesserte das Sprachverstehen hingegen nur wenig. Die kombinierte Änderung von F<sub>0</sub> und F<sub>n</sub> ergab jedoch einen „super-additiven“ Effekt, der die Ergebnisse übertraf, die durch die einzelnen Änderungen von F<sub>0</sub> und F<sub>n</sub> hervorgerufen wurden. Anlehnend an diese Ergebnisse könnte vor allem die bessere Übertragung der F<sub>0</sub> durch das HG das Sprachverstehen von CI-Trägern bei konkurrierenden Sprechern verbessern.

Ziel ist es zu untersuchen, ob CI-Träger mit Hilfe akustischer Reizung des kontralateralen Ohres den Zielsatz besser verstehen als mit dem CI alleine. Es wird weiterhin untersucht, inwieweit ein HG auf dem nicht-implantierten Ohr zu einer besseren Trennung von Ziel- und Maskierersprecher führt. Hinsichtlich der eingeschränkten spektralen Auflösung im CI wird grundlegend erwartet, dass die Wahrnehmung der F<sub>n</sub> nur schlecht gelingt und dadurch ein Rückschluss auf den jeweiligen Sprecher nur bedingt möglich ist. Dies kann dazu führen, dass CI-Träger besonders in störschallerfüllten Kommunikationssituationen nicht in der Lage sind, der Stimme eines Sprechers zu folgen, und dass das Sprachverstehen vermindert wird. Anlehnend an die Studie von Darwin et al. (2003) mit NH wird Hypothese 4 aufgestellt:

*H4: Grundfrequenzunterschiede können besser genutzt werden als Formantfrequenzunterschiede.*

Hinsichtlich des Zugewinns durch das HG wird Folgendes erwartet:

*H5: Die Sprachverständlichkeit bei konkurrierenden Sprechern ist in der bimodalen Kondition besser als für jedes versorgte Ohr allein.*

Dies sollte unabhängig für alle Ziel-Maskierer-Konditionen (Veränderung von F<sub>0</sub> und / oder F<sub>n</sub>) gelten.

*H6: Das HG führt zu einer besseren Trennung der Sprecher, wodurch Verwechslungen von Ziel- und Maskiererwörtern minimiert werden (Brungart, 2001) und IM reduziert wird. Gleichzeitig wird durch die zweite Hörhilfe insgesamt mehr verstanden, so dass auch EM reduziert wird.*

Dies wird der komplementären Information zugeschrieben, die durch die beiden unterschiedlichen Hörhilfen übertragen wird. Der tieffrequente Bereich, der durch das HG übertragen wird, enthält Merkmale, die für die Erkennung verschiedener Stimmen essentiell sind. Der hochfrequente Bereich, übertragen durch das CI, wird durch den tieffrequenten Bereich ergänzt (Ching et al., 2007).

### **6.4 Experiment 3: Effekte räumlichen Hörens**

Das letzte Experiment beschäftigt sich mit der räumlichen Trennung von Ziel- und Maskierersprecher. Zusätzlich zu Unterschieden in F0 zwischen Ziel- und Maskiererstimme, wurde der Maskierersprecher aus unterschiedlichen Richtungen (0°, +/- 90° Azimut-Winkel)<sup>4</sup> präsentiert. Dieses Experiment wurde ebenfalls mit bimodal versorgten Probanden durchgeführt.

Es wird untersucht, ob die von dem CI und dem HG aufgenommenen Signale zentral miteinander verbunden werden und Vorteile des binauralen Hörens genutzt werden können. Dafür werden der Kopfschatteneffekt, der Squelcheffekt und der Summationseffekt berechnet.

Um sich der Frage nach dem Anteil von IM und EM bei Maskierersignalen zu nähern, wird zusätzlich zu dem Sprachmaskierer (IM) ein modulierter Rauschmaskierer (EM) präsentiert. Dieser Rauschmaskierer besitzt das gleiche Langzeitspektrum wie Sprache und die gleiche Struktur der Amplitudeneinhüllenden der jeweiligen Maskierersätze.

Es werden folgende Hypothesen aufgestellt:

*H7: Das Sprachverstehen mit beiden Hörhilfen ist besser als mit einer Hörhilfe alleine und es ergeben sich positive Effekte durch das binaurale Hören.*

Dies sollte für beide Maskiererarten gelten. Bei Verwendung des Rauschmaskierers wird nur EM reduziert. Im Vergleich dazu könnte sich das Sprachverstehen bei Verwendung des Sprachmaskierers durch eine zusätzliche Reduzierung des IM verbessern.

*H8: Nicht nur der Effekt des EM, sondern auch der Effekt des IM wird bei Sprachmaskierern verringert, wenn die Sprecher räumlich getrennt sind.*

---

<sup>4</sup> Ein Winkel von -90° entspricht einem Winkel von 270°

## **Kapitel 6. Zielsetzung und Hypothesen**

---

Ein Unterschied in F0 sollte keinen Einfluss haben, wenn Ziel- und Maskierersignale aus unterschiedlichen Richtungen dargeboten werden, wie in Studien mit NH gezeigt werden konnte (Hawley et al., 2004).

*H9: Die räumliche Trennung ist die dominierende Größe verglichen mit Unterschieden in F0.*

## KAPITEL 7

### Vortest - Sprecherunterscheidung und Sprechererkennung

Für NH gilt, dass  $F_0$  und  $F_n$  wichtige Merkmale sind, durch die Sprecherstimmen identifiziert und voneinander unterschieden werden können.  $F_0$  entspricht der wahrgenommenen Stimmhöhe. Die  $F_n$ , die u.a. durch die Vokaltraktlänge (VTL) definiert werden, beeinflussen den Stimmklang (Darwin et al., 2003). Die Basis der Sprechererkennung ist die Identifikation und Diskrimination des Sprechergeschlechts.  $F_0$  gilt als das wichtigste Merkmal, anhand dessen das Sprechergeschlecht erkannt werden kann.

Den Aussagen von Darwin et al. (2003) zu Folge ist es für Normalhörende einfacher, Änderungen in  $F_0$  wahrzunehmen als Änderungen der VTL. Sie untersuchten die Fähigkeit von normalhörenden Personen, künstliche Unterschiede in der Grundfrequenz und / oder in den Formantfrequenzen zwischen zwei simultanen Sprechern wahrzunehmen. Für NH war  $F_0$  hierbei der dominierende Parameter. Bereits kleine Unterschiede zwischen den Sprechern konnten genutzt werden, um den Zielsprecher besser zu verstehen.

Aufgrund der eingeschränkten Verarbeitung spektraler Merkmale in CIs, könnte die Unterscheidung von Stimmen für CI-Träger ein Problem darstellen.

Die meisten Studien, die sich mit der *Sprecherunterscheidung* bei CI-Trägern beschäftigen, untersuchten nicht gezielt ausgewählte akustische Merkmale. Es wurden unterschiedliche Sprecher genutzt, so dass sich die Stimuli in einer Vielzahl von Merkmalen wie Sprechgeschwindigkeit oder Artikulationsweise unterscheiden, die über  $F_0$  und  $F_n$  hinausgehen. Untersuchungen, in denen künstliche Stimuli verwendet wurden, konzentrierten sich auf die Verarbeitung von  $F_0$ . Um Sprecher des gleichen Geschlechts unterscheiden zu können, ist die Wahrnehmung der sprecherspezifischen  $F_n$  jedoch unerlässlich.



### 7.1 Zielsetzung

In diesem Vortest wurden  $F_0$  und VTL manipuliert und die Diskrimination von Sprecherstimmen und die Identifikation des Sprechergeschlechts bei CI-Trägern getestet.

Inwieweit CI-Träger  $F_n$  nutzen, um nicht nur das Sprechergeschlecht, sondern auch unterschiedliche Sprecher desselben Geschlechts unterscheiden zu können, soll hier untersucht werden.

Anlehnend an eine Studie von Fu et al. (2005), die die *Sprechererkennung* bei CI-Trägern untersuchten, wird angenommen, dass aufgrund der durch das CI übertragenen eingeschränkten spektralen Merkmale  $F_0$  sowohl bei der Sprecherunterscheidung als auch bei der Erkennung des Sprechergeschlechts im Vergleich zu Unterschieden in  $F_n$  auch bei CI-Trägern der dominierende Faktor ist.

### 7.2 Methode und Durchführung

Bei den Stimuli handelt es sich um den Satz „Die Krankenschwester isst einen Apfel“. Dieser wurde aus der Prosodie-Testbatterie (Meister et al., 2008) entlehnt. Der männliche Originalsprecher hat eine, über den gesamten Satz gemittelte,  $F_0$  von 157 Hz und eine Vokaltraktlänge von ca.13,2 cm.

Die Sprecherstimme wurde in  $F_0$  sowohl angehoben als auch abgesenkt. Die  $F_n$  wurden durch eine variierende VTL verändert. Die Stimuli wurden auf der Grundlage der Vorgehensweise von Darwin et al. (2003) generiert. Die männliche Stimme wurde für  $F_0$  mit dem Faktor 1.7 multipliziert und für VTL mit dem Faktor 0.84. In dieser Weise erhält man eine weibliche Stimme, bei der der natürliche Unterschied zwischen männlichen und weiblichen Stimmen eingehalten wird (Peterson und Barney, 1952). Zwischen dem Originalstimulus und der generierten weiblichen Stimme wurden die Faktoren in Darwin et al. (2003) interpoliert, um kleinere Stufen in  $F_0$  und VTL produzieren zu können. Für die vorliegende Studie wurde die Interpolation der Stimuli auf einen größeren Bereich ausgeweitet. Dafür wurden anhand der Werte von Darwin et al. (2003) polynomische Formeln 3. Grades sowohl für  $F_0$  (Formel B.1) als auch für VTL (Formel B.2) angenähert. Die dazugehörigen Werte und Formeln zur Stimuluserstellung befinden sich in Anhang B.1 mit den Tabellen B.1 und B.2 und den Abbildungen B.1 und B.2. Die Manipulationen wurden mithilfe des Sprachanalyseprogramms „PRAAT©“ (Boersma und Weenink, 1996) durchgeführt. Die Basis bildete das Praatskript „change gender“ (s. Anhang B.2) von Darwin et al. (2003).

The apparent vocal-tract length of each talker was changed by a factor of  $vt$  for each utterance by (1) multiplying  $F_0$  by  $vt$  and duration by  $1/vt$  (using PSOLA), (2) resampling at the original

## Kapitel 7. Vortest - Sprecherunterscheidung und Sprechererkennung

sampling frequency multiplied by  $vt$ , and then (3) playing the samples at the original sampling frequency. The end effect of this manipulation was to maintain the same duration and  $F_0$  of the original utterance but to scale the spectral envelope by  $vt$  (Darwin et al., 2003: 2918).

In Experiment 1 zur *Sprecherunterscheidung* wurde nur die Phrase „einen Apfel“, die aus dem oben genannten Satz entnommen wurde, präsentiert. Hierbei wurden nur  $F_0$  oder VTL verändert, um gezielt den Einfluss der einzelnen Parameter getrennt auf die Sprecherunterscheidung untersuchen zu können. Der kleinste Unterschied in  $F_0$  belief sich auf 14 Hz, wohingegen der größte Unterschied in  $F_0$  137 Hz war. Der größte Unterschied in VTL belief sich auf 20%<sup>5</sup> (s. Tabelle B.2 in Anhang B.1).

Auf Grundlage dieser Stimuli wurde mit einem „3-alternative forced choice“ (3-AFC) Verfahren ein Experiment zur Sprecherunterscheidung durchgeführt. D.h. es wurden jeweils drei Stimuli präsentiert, wobei sich ein Stimulus von den übrigen zwei Stimuli unterschied (Referenz). Die Probanden hatten die Aufgabe, den unterschiedlichen Stimulus anzuzeigen. Bei richtigen Antworten wurden die Kategorien in Schritten von 0,5 verringert, bei falschen Antworten wurden die Kategorien um 3 x 0,5 Schritte erhöht. Die Ergebnisse zeigen an, dass die Probanden zu 75%-iger Sicherheit den Unterschied zwischen den Stimuli wahrnehmen konnten. Die Ratewahrscheinlichkeit betrug 33%.

In Experiment 2 wurde die Fähigkeit untersucht, Sprecher zu *erkennen*. Dafür wurde ein „2-alternative forced choice“ (2-AFC) Verfahren verwendet. Es wurde der gesamte o.g. Satz vorgespielt. Die Probanden mussten entscheiden, ob dieser von einem männlichen Sprecher oder einer Sprecherin geäußert wurde. Die Stimme des Originalsprechers wurde in acht Stufen verändert und in einem Bereich von -2 bis +2 in Kategorien von 0,5 kodiert. Hierbei wurden  $F_0$ , VTL und die Kombination  $F_0 + VTL$  getestet. Wie aus Tab 7.1 ersichtlich, kann jeder Kategorie eine entsprechende akustische Änderung in  $F_0$  und VTL bzw.  $F_n$  zugeordnet werden. Die Ratewahrscheinlichkeit betrug 50%. Die Stimuli der unterschiedlichen Kategorien wurden randomisiert dargeboten.

	Weibliche Stimme				Original	Männliche Stimme			
	2	1,5	1	0,5	0	-0,5	-1	-1,5	-2,0
F0	212	199	176	174	157	143	133	121	116
F1	872	863	839	831	807	783	775	751	742
F2	1974	1956	1901	1883	1828	1773	1755	1700	1682
F3	3096	3068	2982	2953	2867	2781	2752	2666	2638

Tabelle 7.1: Änderungen in  $F_0$  und in VTL bzw.  $F_n$  ( $F_1$ ,  $F_2$ ,  $F_3$ ) in Hz für insgesamt 9 Stimuli für Experiment 2: Sprechererkennung

<sup>5</sup> Eine Änderung in VTL entspricht einer Änderung der Formantfrequenzen ( $F_n$ ). Im weiteren Verlauf der Arbeit werden zur Vereinfachung die Begriffe VTL und  $F_n$  gleichgesetzt.

Die Messungen fanden in einem schallarmen Raum statt. Die Signale wurden über einen Lautsprecher, der in einem Abstand von 1,2 m frontal vor dem Patienten stand, bei 70 dB SPL dargeboten. Zu Übungszwecken wurden die extremsten Kategorien und der Originalstimulus vorgespielt und Feedback über die Korrektheit der Antworten durch den Untersucher gegeben. Bei den eigentlichen Messungen erfolgten die Antworten der Patienten über einen Touch-Screen.

An der Vormessung nahmen 6 unilateral versorgte CI-Probanden teil, mit denen später auch Experiment 1 durchgeführt wurde. Nähere Angaben über Demografie der CI-Probanden finden sich in Tabelle B.5.1 im Anhang B.5. Eine Gruppe von 6 NH diente als Kontrollgruppe. Die Hörschwelle der NH lag in den Prüffrequenzen zwischen 250 Hz und 6 kHz über 20 dB HL. Die NH waren zwischen 23 und 38 Jahren alt ( $\bar{\varnothing}$  28 Jahre).

### **7.3 Ergebnisse**

In diesem Unterkapitel werden die Ergebnisse der zwei Experimente vorgestellt. Da dieser Vortest jedoch nur als Orientierung für die folgenden Experimente galt, erfolgt die Ergebnissanalyse rein deskriptiv ohne Berechnung der statistischen Signifikanzen.

#### **7.3.1 Experiment 1: Sprecherunterscheidung**

Abbildung 7.1 zeigt die Ergebnisse für das Diskriminationsexperiment. Dargestellt ist der Unterschied zwischen den Sprechern in Prozent, der von den Probanden wahrgenommen wurde. Die Mittelwerte der 6 CI-Probanden und der NH-Kontrollgruppe sind für den Parameter  $F_0$  und  $F_n$  aufgetragen.

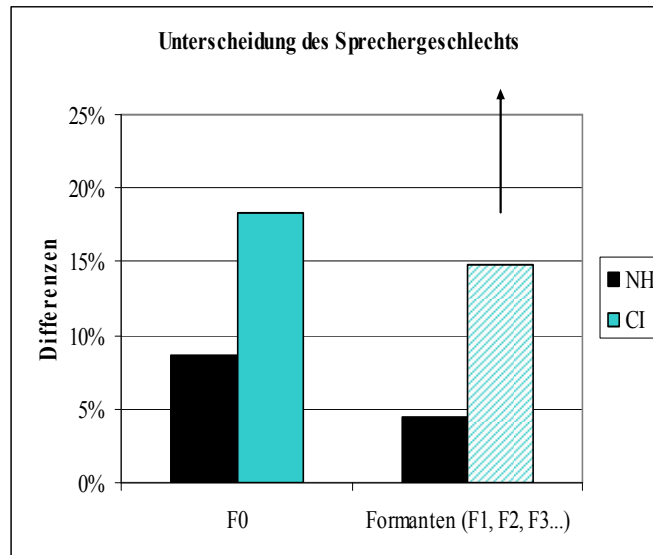


Abbildung 7.1: Ergebnisse für die Unterscheidung des Sprechergeschlechts; dargestellt als eben wahrnehmbare Schwelle (just-noticeable difference, JND) in % für F0 (links) und für die Formantfrequenzen (rechts) basierend auf Änderung der VTL.

Werte des Originalsprechers: F0: 157 Hz; F1: 807 Hz; F2: 1828 Hz; F3: 2867 Hz

NH waren in der Lage einen Unterschied von 8% in F0 zu diskriminieren. Der Vergleich zu den NH zeigt, dass CI-Nutzer einen mehr als doppelt so hohen F0-Unterschied von 18% benötigten, um Sprecher unterscheiden zu können.

Die Ergebnisse der Veränderung der VTL und somit der  $F_n$  sind nicht eindeutig. Das liegt daran, dass mehr als die Hälfte der CI-Probanden viel größere Unterschiede in VTL gebraucht hätte, als für diesen Test generiert wurden. Der größte Unterschied belief sich auf Änderungen um 20% im VTL verglichen zu der Originalstimme. Die Diskriminationsschwelle der CI-Nutzer gingen über diese 20%ige Änderung hinaus. Bei den NH lag die Diskriminationsschwelle bei ca. 5%.

### 7.3.2 Experiment 2: Sprechererkennung

Abbildungen 7.2.1, 7.2.2 und 7.2.3 zeigen die gemittelten Ergebnisse für die 6 CI-Träger und die NH-Kontrollgruppe für Änderungen in F0 (s. Abbildung 7.2.1), in VTL (s. Abbildung 7.2.2) und in der Kombination F0 + VTL (s. Abbildung 7.2.3). Auf der x-Achse sind die Messpunkte der neun Stufen (-2 bis +2) aufgetragen. Gezeigt wird die Wahrscheinlichkeit  $p$ , dass die weibliche Stimme als weiblich erkannt wurde, so dass der negative Bereich in der Darstellung der Antwort „nicht weiblich“ bzw. „männlich“ entspricht. Dargestellt sind die Daten anhand einer Diskriminationskurve, deren Steigung anzeigt, ob Unterschiede diskriminiert werden konnten. Je steiler die Kurve, desto sicherer gelang die Zuordnung der Kategorien zu einer weiblichen bzw. männlichen Stimme. Je flacher der Verlauf, desto

unsicherer war die Zuordnung. Die Verschiebung der Kurve auf der x-Achse wird nicht betrachtet, da diese „eher Hörgewohnheiten als Perzeptionsleistungen widerspiegelt.“ (Meister et al., 2008: 344). Insgesamt zeigt sich, dass der Originalsprecher von allen NH als männlich erkannt wurde ( $F_0 = 157$  Hz). Bei den CI-Trägern wurde diese Stimme weder als eindeutig männlich, noch als eindeutig weiblich wahrgenommen.

Der steile Anstieg, der bei Änderungen in  $F_0$  (s. Abbildung 7.2.1) auftrat, macht deutlich, dass dieser Parameter gut genutzt werden konnte, um das Sprechergeschlecht zu identifizieren. Jede Änderung der Originalstimme konnte als männlich oder weiblich identifiziert werden.

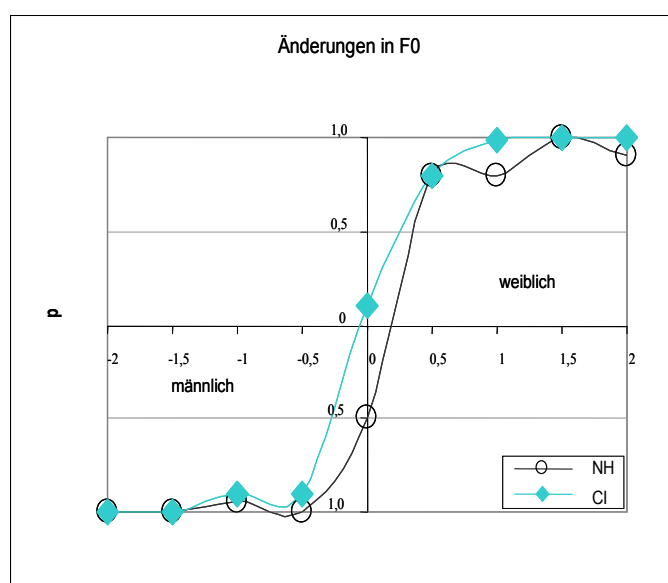


Abbildung 7.2.1: Ergebnisse für die Identifikation des Sprechergeschlechts. Änderungen in  $F_0$ ; (Ratewahrscheinlichkeit  $p = 0,5$ ).

Die Ergebnisse für Änderungen in der VTL (s. Abbildung 7.2.2) zeigt, dass weder NH noch CI-Probanden das Sprechergeschlecht anhand dieses Parameters eindeutig identifizieren konnten. Die Ergebnisse für Änderungen in der VTL lagen eher auf Niveau der Ratewahrscheinlichkeit ( $p = 0,5$ ).

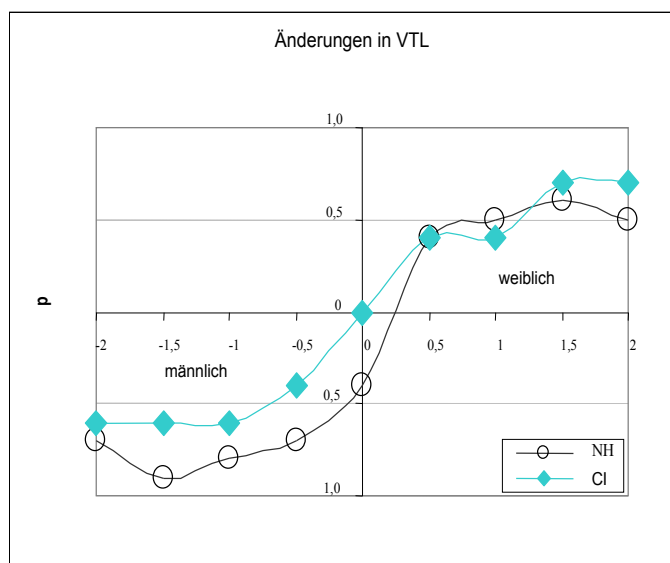


Abbildung 7.2.2: Ergebnisse für die Identifikation des Sprechergeschlechts. Änderungen in VTL; (Ratewahrscheinlichkeit  $p = 0.5$ ).

Durch Änderungen in der Kombination F0 + VTL (s. Abbildung 7.2.3) verbesserte sich die Sprechergeschlechterkennung in der NH-Kontrollgruppe zusätzlich im Vergleich zu den Änderungen nur in F0 (s. Abbildung 7.2.1). Die Diskriminationskurve verläuft etwas steiler und nahezu jeder Stimulus wurde eindeutig als männliche oder weibliche Stimme identifiziert. Dieser leicht additive Effekt von F0 und VTL ergab sich bei den CI-Probanden nicht. Die Sprechergeschlechterkennung bei Änderungen von F0 + VTL entsprach in etwa den Ergebnissen der Änderungen nur von F0.

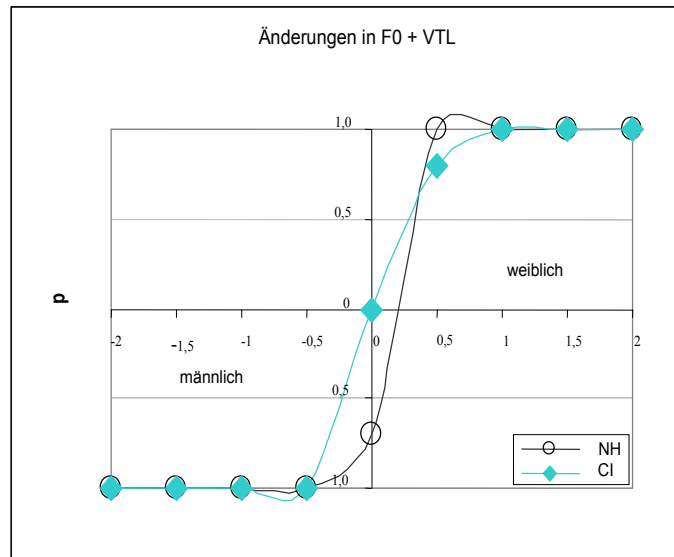


Abbildung 7.2.3: Ergebnisse für die Identifikation des Sprechergeschlechts. Änderungen in F0 und VTL; (Ratewahrscheinlichkeit  $p = 0.5$ ).

### 7.4 Diskussion

Das Ziel dieser Messungen war herauszufinden, inwieweit unilateral versorgte CI-Träger in der Lage sind, unterschiedliche Stimmen zu diskriminieren und zu identifizieren. Es galt definierte Änderungen der akustischen Merkmale Grundfrequenz und der Formantfrequenzen wahrzunehmen.

Im Unterscheidungsexperiment zeigte sich, dass CI-Träger durchaus in der Lage waren, verschiedene Stimmen anhand von Manipulationen in F0 zu unterscheiden, wobei die Diskriminationsschwelle 2 - 3 mal so hoch lag wie die der NH. Die Diskriminationsschwelle für die Formantfrequenzen war im Durchschnitt mehr als 4-mal so hoch. Viele der CI-Probanden konnten den größtmöglichen Unterschied in den präsentierten Stimuli nicht unterscheiden.

Die Ergebnisse entsprechen denen des Identifikationsexperiments. Beide Probandengruppen waren in der Lage, männliche und weibliche Stimmen zu identifizieren, wenn entweder die Parameter F0 alleine oder die Kombination F0 + VTL manipuliert wurden. Hierbei zeigten sich keine großen Unterschiede zwischen der normalhörenden Kontrollgruppe und den CI-Trägern. Änderungen der VTL alleine (d.h. der Formantfrequenzen) reichten hingegen nicht aus, um das Sprechergeschlecht zu identifizieren. Die zu Änderungen in F0 hinzugefügten Änderungen der VTL verbesserten die Fähigkeit der Normalhörenden das Sprechergeschlecht zu erkennen. Die CI-Träger konnten jedoch nicht von der zusätzlichen Information profitieren.

Vermutlich stützt sich die Fähigkeit der CI-Träger, Sprecher zu identifizieren, eher auf Informationen der Periodizität, repräsentiert durch Änderungen in F0, als auf die Auswertung

spektraler Merkmale. Dadurch können größere Schwierigkeiten auftreten, wenn es darum geht, Sprecherstimmen des gleichen Geschlechts zu unterscheiden. Hierbei spielt die Wahrnehmung von Unterschieden in den Formantfrequenzen eine wichtigere Rolle als bei der Erkennung und Unterscheidung des Sprechergeschlechts. Dafür spricht auch eine Studie von Vongphoe und Zeng (2005), die zeigen konnten, dass die Erkennung des Sprechergeschlechts grundsätzlich möglich ist für CI-Träger. Die Sprechererkennung bei Personen des gleichen Geschlechts, bei der die Formantfrequenzen eine größere Rolle spielen, bleibt jedoch problematisch.

### **7.5 Fazit**

CI-Träger können mit wenigen Problemen anhand von Modifikationen der Grundfrequenz Sprecher unterscheiden und das Sprechergeschlecht identifizieren. Hingegen gelingen die Unterscheidung von Stimmen und die Identifikation des Sprechergeschlechts auf Grundlage von Modifikationen der Formantfrequenzen nicht. Durch die Modifikation beider Parameter gleichzeitig ergibt sich kein additiver Effekt. Dies ist unterschiedlich zu den Ergebnissen der Normalhörenden. Ihnen gelingt die Identifikation des Sprechergeschlechts auf Basis von Formantunterschieden zwar ebenfalls nicht, es ist ihnen jedoch möglich, diese Merkmale bei der Unterscheidung von Sprecherstimmen zu nutzen. Bei kombinierter Änderung der Parameter zeigt sich ein additiver Effekt.

In den folgenden Experimenten stellt sich nun die Frage, ob Grundfrequenzunterschiede von CI-Trägern genutzt werden, um synchrone Sprecher zu differenzieren. Da die Formantunterschiede durch das CI nicht ausgewertet werden konnten, wird dieses Merkmal in Experiment 2 gezielt mit Hilfe eines zusätzlichen Hörgerätes untersucht, in der Annahme, dass die Feinstruktur der Sprachsignale bei akustischer Verstärkung noch erhalten bleibt und übertragen werden kann.



## Kapitel 8

### Methode und Durchführung

Im deutschsprachigen Raum stand bis jetzt kein sprachaudiometrischer Test zur Verfügung, mit dem das Phänomen IM untersucht werden konnte. Im Folgenden wird der von Wagener et al. (1999a-c) entwickelte Oldenburger Satztest (OISa) vorgestellt, der in der klinischen Routine als Messinstrument eingesetzt wird. Darauf aufbauend wird erläutert, wie dieser Satztest für IM-Messungen genutzt werden kann.

#### 8.1 Oldenburger Satztest

Der OISa wird häufig als Diagnostikinstrument bei der Versorgung und Kontrolle von Hörhilfen eingesetzt. Mithilfe des Tests kann das Sprachverstehen von Patienten in Ruhe und im Störgeräusch beurteilt werden.

Beim OISa handelt sich um einen quasi-geschlossenen Test, bei dem das Sprachmaterial aus kontextfreien „5 Wort“-Sätzen besteht. Die Struktur der Sätze ist immer die gleiche und besteht aus Vorname-Verb-Zahl-Adjektiv-Nomen. Ein Beispielsatz ist „Stefan verleiht achtzehn schwere Blumen“. Für jede Wortgruppe stehen 10 Auswahlmöglichkeiten zur Verfügung. Es existiert eine Basisliste mit 10 Sätzen, die der Phonemverteilung des Deutschen entsprechen (s. Anhang B.3). Es wurden nicht alle möglichen ( $10^5$ ) Kombinationen aufgesprochen, sondern 100 Sätze erstellt, „bei denen jeder Übergang von einem bestimmten Wort zu jedem der zehn möglichen nachfolgenden Wörter genau einmal vorkommt“ (Wagener et al., 1999a: 11). Somit wurden Koartikulationseffekte beibehalten und ein Kompromiss zwischen Natürlichkeit und zu vertretendem Arbeitsaufwand gefunden. Durch randomisierte Kombination der Wörter gelingt es, die Vielzahl von  $10^5$  verschiedenen Sätzen zu kreieren. Zusätzliche Merkmale wie z.B. Prosodie, die ein sog. „Streaming“ erlauben, gibt es allerdings nicht. Das bedeutet, dass nur wenig Vorabinformation vorhanden ist, durch die die Aufeinanderfolge der Wörter vorhergesagt werden könnte. Durch die kontextfreie Gestaltung des OISa ist es den Hörern ebenfalls nicht möglich, sich das nächste

Wort zu erschließen. Aufgrund der niedrigen Vorhersagbarkeit des OISa, können Testlisten somit mehrmals verwendet werden (Wagener et al., 1999c).

Es gibt eine Lernphase, in der die Patienten ein Wissen über das vorhandene Vokabular des OISa erlangen. In der klinischen Routine zeigt sich, dass das Sprachverstehen nach einer Trainingsphase mit zwei Testlisten in etwa konstant bleibt. Der OISa kann somit als stabiles Messinstrument genutzt werden (Wagener et al., 1999c).

Der OISa wird von einem männlichen Sprecher mit einer F0 von durchschnittlich 120 Hz in mittlerer bis gemäßiger Sprechgeschwindigkeit (233 Silben/Minute) gesprochen (Wagener et al., 1999a).

### **8.2 Oldenburger Satztest bei IM**

Anlehnend an die Struktur des CRM-Tests (vgl. Kapitel 4) ist es möglich, Messungen zum IM mithilfe des OISa (Wagener et al., 1999 a-c) durchzuführen. Dieser eignete sich aus verschiedenen Gründen:

- 1) Aufgrund der besonderen Struktur des OISa („Stefan sieht acht nasse Sessel“) kann ein Schlüsselwort definiert werden (z.B. der Name), mit dem der Zielsatz eindeutig gekennzeichnet werden kann.
- 2) Bei den Aufnahmen fanden zwar Koartikulationseffekte Berücksichtigung, Merkmale wie z.B. Prosodie kommen jedoch nicht vor. Dadurch ist der Verlauf der Sätze durch Streamingmechanismen nicht möglich und eine Unterscheidung zwischen Ziel- und Maskiersatz erschwert.
- 3) Die Testlisten sind homogen, da sie alle aus demselben Wortmaterial bestehen (s. Anhang B.3).
- 4) Es ist nicht möglich, die Testlisten während einer Untersuchung auswendig zu lernen. Gleiche Testlisten können somit mehrmals bei denselben Probanden verwendet werden.

Wie in Pyschny et al. (2009) beschrieben, wurden für die Experimente zum IM, Sätze gleicher Länge überlagert, um eine möglichst synchrone Maskierung zu erhalten. Das gewählte Schlüsselwort, das die jeweiligen Zielsätze kennzeichnet, war in allen Experimenten der Vorname „Stefan“. Im Rahmen von Voruntersuchungen mit Normalhörenden zeigte sich, dass dieser Vorname das geringste Risiko barg, mit den anderen Vornamen des OISa verwechselt zu werden. Außerdem lag der Name „Stefan“ im mittleren Bereich der Verständlichkeit, was bedeutet, dass dieser Name weder zu schlecht wahrnehmbar, noch zu dominant war. Die Maskiersätze beinhalteten alle restlichen

Namen. Bei der Überlagerung von Ziel- und Maskierersätzen unterschieden sich die Wörter der Zielsätze immer von denen der Maskierersätze.

Die Sätze maskieren unterschiedlich stark, weil es in Bezug auf die Pegel der Sätze leichte Schwankungen von durchschnittlich ca. 2 dB gab. Dies hatte jedoch über die späteren Testlisten keinen Effekt, so dass diese natürlichen Schwankungen beibehalten wurden.

Nach Brungart (2001) kann IM in maximaler Ausprägung durch maximale „target-masker similarity“ und damit maximale „stimulus uncertainty“ erreicht werden. Dies erreicht man durch identische Ziel- und Maskierersprecher. Unterschiedliche  $F_0$ , unterschiedliche  $F_n$  und andere stimmliche Variablen können als Merkmale vom Hörer genutzt werden, um Sprecher zu trennen und somit IM zu reduzieren. Das theoretische Konstrukt der gleichen Ziel- und Maskiererstimme, mit dem man versucht, möglichst viel IM zu gewährleisten, führt dazu, dass NH zwar in der Lage sind, alle Wörter der konkurrierenden Sprecher zu verstehen, die Wörter aber nicht dem jeweiligen Sprecher zuordnen können. Es werden 50% der Wörter des Zielsatzes und 50% der Wörter des Maskierersatzes verstanden. Die Fehler, die hierbei gemacht werden, sind somit reine Verwechslungen zwischen Ziel- und Maskiererwörtern. Nach Brungart (2001) stehen solche Fehler für IM. Werden Wörter falsch oder gar nicht verstanden, würde dies auf das Phänomen EM zurückzuführen sein.

Es wurden insgesamt 60 Satzpaare generiert, die in 4 Testlisten á 15 Satzpaare aufgeteilt wurden. Jede der 4 Testlisten erfüllte das Kriterium der Listenäquivalenz und die Antwortverteilung von 50% Ziel- und 50% Maskiererwörtern. Vorversuche mit NH zeigten, dass es keine Listeneffekte oder systematische Fehler gab. In allen Experimenten wurden diese Testlisten verwendet. Die Überlagerung wurde mithilfe eines Matlab-Skripts automatisiert.

Das so generierte Inventar des OISa wurde in den drei folgenden Experimenten als Messinstrument für die Untersuchungen zum IM genutzt.

Von denen in Kapitel 4 beschriebenen Faktoren, die die Trennung konkurrierender Sprecher beeinflussen, werden insgesamt in den Experimenten der TMR, die stimmliche Charakteristik, die räumliche Trennung und bei allen Experimenten das Wissen über Identität und Ort des Sprechers berücksichtigt. Die Anzahl der Maskierersprecher zu erhöhen, würde den Anteil an IM erhöhen (Cullington und Zeng, 2008). Allerdings können mit nur einem Maskierersprecher einerseits die Effekte von unterschiedlichen stimmlichen Charakteristiken und andererseits die Fehlerverteilung detaillierter untersucht werden.

### **8.3 Stimuli der Maskierungsexperimente**

Sprachaudiometrische Tests werden im deutschsprachigen Raum allesamt von einer männlichen Stimme gesprochen. Anlehnend an diese Gegebenheit wurde auch in den folgenden Maskierungsexperimenten mit der männlichen Zielstimme gearbeitet. Die Maskierersätze wurden aus der Zielstimme generiert. Die Maskiererstimme wurde künstlich manipuliert und war in allen Experimenten entweder von gleicher oder höherer F0 als die Zielstimme.

Die Original-Sprecherstimme des OISa wurde mithilfe des Analysetools PRAAT© (Boersma und Weenink, 1996) um 20 Hz verringert und lag im Durchschnitt bei 100 Hz. Dies hatte den Vorteil, dass alle Stimuli mit Praat bearbeitet wurden, so dass eventuelle Artefakte aufgrund der Signalverarbeitung durch Praat in allen Signalen enthalten waren.

Weiterhin wurde die Stimme herabgesetzt, um bei Veränderung der Maskiererstimme, eine größere Spannbreite zu höheren F0 zur Verfügung zu haben. Zu große Änderungen in F0 führen zu Verzerrungen und Artefakten, die die Stimme unnatürlich klingen lassen. Auch wenn sich die Wahrnehmung der Stimme bei Übertragung mit einem CI nicht mit der Wahrnehmung über ein gesundes Ohr vergleichen lässt, kann eine unangenehm empfundene Sprechtonhöhe den nachfolgende Wahrnehmungs- und Wirkungsprozess beeinträchtigen (Eckert und Laver, 1994).

Es wurden keine unterschiedlichen natürlichen Sprecher verwendet, sondern explizit Stimuli mit einer manipulierten Maskiererstimme generiert, um gezielt den Effekt stimmlicher Parameter und deren Auswirkung auf IM untersuchen zu können.

#### **8.3.1 Experiment 1: Pegel und Grundfrequenzunterschiede**

In Experiment 1 wurden die Maskierersätze hinsichtlich der Grundfrequenz (F0) und des Target-to-masker ratios (TMR) verändert. F0 wurde von 100 Hz bis 180 Hz in 20 Hz-Schritten erhöht. Die Einteilung wurde nach Trinh et al. (2006) vorgenommen:

100 Hz Zielsatz - 100 Hz Maskierer	= same talker
100 Hz Zielsatz - 120 Hz oder 140 Hz Maskierer	= same sex
100 Hz Zielsatz - 160 Hz oder 180 Hz Maskierer	= different sex

Dafür wurde das in Praat implementierte „change gender“ Tool genutzt, das mit dem PSOLA-Algorithmus (engl. “pitch synchronous overlap and add“) arbeitet. Änderungen der Amplitude oder Änderungen in zeitlichen Merkmalen gibt es dabei nicht (Boersma und Weenink, 1996).

## Kapitel 8. Methode und Durchführung

---

Der TMR erstreckt sich von 0 dB bis 20 dB, mit einer Schrittweite von 5 dB. Diese Änderungen wurden mithilfe des Programms „Cool Edit“ vorgenommen. Das Vorkommen von Pegelunterschieden zwischen Ziel- und Maskierersätzen bedeutet, dass die Probanden ein zusätzliches akustisches Merkmal ausnutzen könnten, um die konkurrierenden Sprecher voneinander zu trennen.

### 8.3.2 Experiment 2: Grund- und Formantfrequenzunterschiede

In Experiment 2 wurde ausschließlich die stimmliche Charakteristik untersucht. Es wurden sowohl  $F_0$  als auch  $F_n$  des Maskierers modifiziert. Hierzu wurde ein Praatskript (s. Anhang B.2) verwendet, wie es auch in der Studie von Darwin et al. (2003) eingesetzt wurde und wie es auch hier bereits beim Vortest (Kapitel 7) zum Einsatz kam. Vorwegnehmend kann gesagt werden, dass aufgrund der Ergebnisse von Experiment 1 (s. Kapitel 10) die Schrittweite von  $F_0$  erhöht wurde. Die Stimme des Maskierers wurde in 40 Hz Schritten in einem Bereich von 100 Hz bis 180 Hz künstlich verändert.  $F_n$  blieb hierbei unverändert.

Im Original hat der Sprecher eine Vokaltraktlänge (VTL) von ca. 20 cm. Um  $F_n$  zu verändern, wurde der Vokaltrakt des Maskierers um 8% bzw. um 16% rechnerisch verkürzt. Somit wurde  $F_n$  erhöht. Hierbei blieb  $F_0$  unverändert. Sowohl durch die Änderungen in  $F_0$  als auch in VTL wurde eine weibliche Stimme erzeugt.

In der Kombination gab es Stimuli mit  $\Delta F_0 + \Delta F_n = 40 \text{ Hz} + 8\%$  und  $\Delta F_0 + \Delta F_n = 80 \text{ Hz} + 16\%$ . Andere Kombinationen wie  $\Delta F_0 + \Delta F_n = 40 \text{ Hz} + 16\%$  und  $\Delta F_0 + \Delta F_n = 80 \text{ Hz} + 8\%$  wurden aufgrund unnatürlicher Grundfrequenz-Formantverhältnisse nicht generiert (vgl. Pyschny et al., 2011).

	<b>F0</b>	<b>VTL</b>	<b>F1</b>	<b>F2</b>	<b>F3</b>
<b>Original</b> F0 & $F_n$ unverändert	<b>100 Hz</b>	<b>19,8 cm</b>	<b>353 Hz</b>	<b>1408 Hz</b>	<b>2553 Hz</b>
$\Delta F_0 + \Delta F_n$ = 40 Hz + 8%	140 Hz	18,3 cm	380 Hz	1529 Hz	2771 Hz
$\Delta F_0 + \Delta F_n$ = 80 Hz + 16%	180 Hz	17 cm	409 Hz	1662 Hz	2961 Hz

Tabelle 8.1: Werte für die Grundfrequenz und die Vokaltraktlänge mit einhergehenden Formantfrequenzen der Maskiererstimuli.

Exemplarisch gemessen am Schwa-Laut [ə] im Wort „nasse“

Auch wenn bei einer Änderung von 8% bereits die Stimmindividualität verloren geht (Darwin et al., 2003), wurde eine Änderung von 16% generiert. Der Unterschied sollte groß genug sein, um von den CI-Probanden wahrgenommen werden zu können.

Pegelunterschiede wurden in Experiment 2 nicht eingeführt, um ausschließlich den Einfluss von  $F_0$  und  $F_n$  untersuchen zu können.

### **8.3.3 Experiment 3: Effekte räumlichen Hörens**

In Experiment 3 wurde auf räumliche Aspekte eingegangen. Die Maskiererstimuli wurden nicht nur aus dem Lautsprecher dargeboten, der direkt vor dem Hörer stand ( $0^\circ$  Azimut-Winkel), sondern ebenfalls aus Lautsprechern, die nach rechts ( $+90^\circ$  Azimut-Winkel) oder links ( $-90^\circ$  Azimut-Winkel) verschoben waren. Daraufhin wurde der Kopfschatten-, der Squelch- und der Summationseffekt berechnet. Die Ergebnisse sollen eine Aussage darüber geben, inwieweit binaurale Prozesse genutzt werden können und das Sprachverstehen verbessert werden kann.

Weiterhin wurden zweierlei Maskierer verwendet, um eine Vergleichbarkeit zu klinisch durchgeführten Messungen zu erhalten. Einerseits wurden Sprachmaskierer mit einer  $F_0$  von 100 Hz und 180 Hz generiert, andererseits Rauschmaskierer. Die Rauschmaskierer entsprachen im Frequenzbereich dem Langzeitspektrum des Satzmaterials und wurden ebenfalls auf der Basis der Sprechergrundfrequenz von 100 Hz und 180 Hz generiert. Um ein Rauschsignal zu erzeugen, welches den Sätzen möglichst ähnlich war, wurde das Rauschen im Zeitbereich mit der Amplitudeneinhüllenden der individuellen Maskierersätze moduliert (Festen und Plomp, 1990). Auf diese Weise konnte ein fluktuierendes Rauschen erzeugt werden. Das Signal wurde dafür gleichgerichtet und mit einem Butterworth-Tiefpassfilter bearbeitet. Die Dämpfung beträgt 3 dB bei der Grenzfrequenz 40 Hz (Festen und Plomp, 1990).

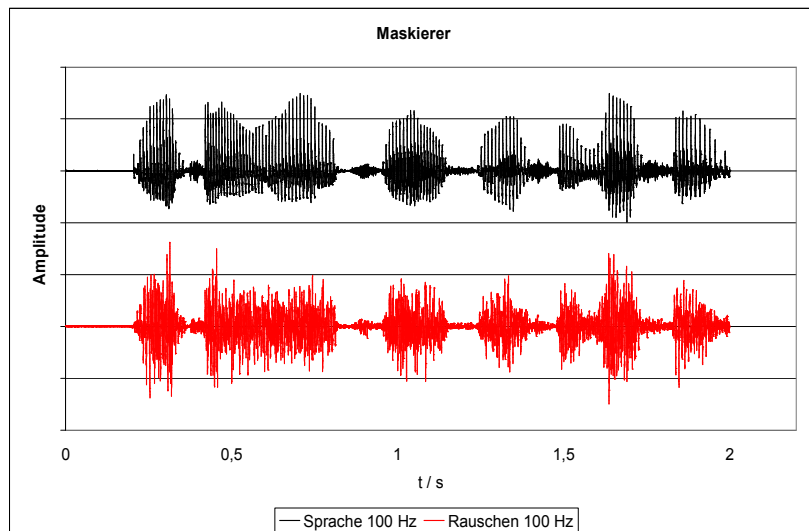


Abbildung 8.1: Darstellung der Maskiererarten im Zeitbereich. Oben: Sprachmaskierer; unten: Rauschmaskierer.

Rausch- und Sprachmaskierer unterschieden sich durch die spektro-temporale Feinstruktur und durch die fehlende Information im Rauschmaskierer. Mithilfe der beiden Maskierer sollte ein Vergleich der Effekte, die durch IM und EM hervorgerufen werden, möglich sein.

Der TMR wurde wie in Experiment 2 in allen Konditionen auf 0 dB festgesetzt.

### 8.4 Probandenkollektiv

Alle Probanden wurden vom Cochlear Implant Centrum der Uniklinik Köln rekrutiert. An allen drei Experimenten nahmen ausschließlich erwachsene CI-Träger teil, die entweder unilateral mit einem CI oder bimodal mit einem CI und einem kontralateralen HG versorgt waren. Das Probandenkollektiv war in den drei Experimenten jedoch unterschiedlich. Die Probanden trugen Systeme der Hersteller Advanced Bionics, Cochlear und Medel. Die Mehrzahl der Probanden war postlingual ertaubt. Die prälingual ertaubten Probanden waren vor der CI-Versorgung mit HGs versorgt und hatten ein gutes bis sehr gutes Sprachverstehen.

Die Experimente wurden in einem zeitlichen Abstand von mehreren Monaten durchgeführt. Aufgrund der im vorherigen Experiment gewonnenen Erfahrungen, wurden die Inklusionskriterien für die Probanden, die an dem jeweils nachfolgenden Experiment teilnahmen, verändert. Gemeinsam war allen drei Experimenten, dass ein Sprachverstehen mit dem CI von mindestens 90% in Ruhe erreicht werden musste. Dies wurde mit dem OISa bei einem Darbietungspegel von 70 dB SPL gemessen. Diese hohen Anforderungen wurden an die Probanden gestellt, um sicherzustellen, dass in Situationen mit Störgeräusch noch genügend Sprache verstanden wurde.

Die bimodal versorgten CI-Träger zeigten auf der HG-Seite durchgehend einen mittel- bis hochgradigen Hörverlust. Tonaudiogramme (unversorgt) und Aufblähkurven (mit HG

versorgt) der nicht implantierten Seite befinden sich, getrennt für Experiment 2 und 3, in Anhang B.6. Für die bimodal versorgten CI-Träger wurde in Experiment 2 ein Sprachverstehen von 50% in Ruhe als Inklusionskriterium festgelegt. Dies wurde mit dem OISa bei einem Darbietungspegel von 70 dB SPL gemessen. Für Experiment 3 wurde dieser Wert auf 75% erhöht, denn es zeigte sich, dass ein Sprachverstehen von 50% in Ruhe nicht ausreichte, um in einer konkurrierenden Sprecher-Situation, Sprache mit dem HG alleine zu verstehen. Das bimodal versorgte Probandenkollektiv ist sehr verschieden von Probandenkollektiven anderer, auswärtiger Studien. Dies ist bei einer Verallgemeinerung der Ergebnisse unbedingt zu berücksichtigen.

Die Messungen in Experiment 1 wurden mit den alltäglichen Sprachprozessoreinstellungen der CI-Nutzer vorgenommen. Das kontralaterale Ohr wurde mit einem Gehörschutz-Kopfhörer (ENHA 3001) verschlossen.

Für die Experimente 2 und 3 wurden in der alltäglichen Sprachprozessoreinstellung Änderungen vorgenommen. Vorverarbeitungsstrategien (z.B. Rauschunterdrückung) und Richtmikrofoncharakteristiken wurden ausgeschaltet, um die Messungen nicht zu verfälschen. Mit eingestellten Vorverarbeitungsmechanismen könnten Störgeräusche insgesamt vermindert werden bzw. aus unterschiedlichen Richtungen kommende Störsignale unterschiedlich stark abgedämpft werden. Dies galt gleichermaßen für die HGs. Vor den Messungen wurde überprüft, ob der subjektive Lautheitseindruck für das CI und das HG vergleichbar waren. Nach Ching et al. (2004) führt eine ähnliche Lautstärkeempfindung beider Hörhilfen zu einem besseren Sprachverstehen. Mithilfe der kategorialen Lautheitsskalierung (Brand und Hohmann, 2002) konnte frequenzabhängig der Zusammenhang zwischen Pegel und Lautheitsempfinden festgestellt werden. Dem rechten und dem linken Ohr wurde getrennt nacheinander sprachsimulierendes Rauschen (Bandbreite 85 Hz bis 5,5 kHz) mit variierendem Pegel über einen Lautsprecher dargeboten. Die Aufgabe der Probanden war, auf einer visuellen Skala, die von „nicht gehört“ über „mittellaut“ bis „zu laut“ reichte, anzugeben, wie laut sie ein Geräusch gehört haben. Die Ergebnisse der Probanden von Experiment 2 zeigten auf, dass die Hörschwellen von CI und HG im Mittel nur knapp 5 dB auseinander lagen. Der maximale Unterschied belief sich auf 10 dB. Auch die Lautheitsanstiegsfunktion war für beide Hörhilfen vergleichbar. Mithilfe der Lautheitsskalierung konnte ebenfalls festgestellt werden, dass die Stimuli für alle Probanden hörbar waren. In Experiment 3 belief sich der Unterschied zwischen den Hörschwellen sogar nur auf 2 dB. Auch hier war die Lautheitsanstiegsfunktion beider Hörhilfen vergleichbar.

Alle Probanden unterschrieben eine Einverständniserklärung, die die anonymisierte Nutzung ihrer Daten erlaubt. Die Probanden erhielten eine Aufwandsentschädigung für jede Sitzung. Detaillierte Angaben über die Demografie der CI-Probanden aller drei Experimente enthalten die Tabellen B.5.1 - B.5.3 in Anhang B.5.



### **8.4.1 Experiment 1**

An den Messungen zu Experiment 1 nahmen sechs unilateral versorgte CI-Träger teil. Es handelte sich um vier weibliche und zwei männliche Hörer im Alter von 39 bis 60 Jahren ( $\bar{x}$  50 Jahre) (Pyschny et al., 2009). Es handelt sich hierbei um dieselben Patienten, die auch an dem Vortest teilgenommen haben.

### **8.4.2 Experiment 2**

An den Messungen zu Experiment 2 nahmen zwölf bimodal versorgte CI-Träger teil. Es handelte sich um sechs weibliche und sechs männliche Hörer im Alter von 18 bis 72 Jahren ( $\bar{x}$  56 Jahre) (Pyschny et al., 2011).

### **8.4.3 Experiment 3**

An den Messungen zu Experiment 3 nahmen zehn bimodal versorgte CI-Träger teil. Es handelte sich um vier weibliche und sechs männliche Hörer im Alter von 19 bis 72 Jahre ( $\bar{x}$  52 Jahre). Hierbei wurde darauf geachtet, dass die Probanden das CI mehr als 9 Monate trugen. Von bilateral versorgten Probanden ist bekannt, dass sich binaurale Fähigkeiten erst über einen längeren Zeitraum entwickeln (Buss et al., 2008).

## **8.5 Durchführung**

Obwohl die Mehrzahl der Patienten bereits aus klinischen Messungen mit dem OISa vertraut war, ging allen Experimenten ein Training voraus. Damit wurde die Kenntnis über das Wortmaterial des OISa und der Zielstimme gefestigt. Dies sollte Lerneffekte für spätere Messungen verringern. Außerdem wurden die Probanden an die Testsituation mit konkurrierenden Sprachsignalen herangeführt.

Zuerst bekamen die Probanden drei Listen des OISa bei 70 dB SPL in Ruhe vorgespielt. Zwei dieser Listen dienten als Übungsmaterial. Bei der dritten präsentierten Liste wurde die Sprachverständlichkeit in %-korrekt festgehalten. Im Anschluss daran wurde mindestens eine Testliste á 20 Sätze in der mit einem Störgeräusch überlagerten Version vorgespielt. In dieser Trainingsphase erhielten die Probanden „Feedback“ darüber, ob ihre Antworten richtig oder falsch waren. Sowohl die Listen als auch die innerhalb der Listen verwendeten Satzpaare wurden den Probanden in randomisierter Reihenfolge angeboten, um Sequenz- und Lerneffekte einzuschränken.

Jedes überlagerte Satzpaar wurde von einem Ton (1 kHz; 500 ms) eingeleitet, um die Aufmerksamkeit der Probanden auf die Stimuli zu lenken. Die Probanden hatten in allen drei Experimenten die Aufgabe, den Satz zu wiederholen, der mit dem Namen „Stefan“ begann. Die Stimuli waren in ein computergestütztes Messprogramm eingebunden, mithilfe dessen der Untersucher die Antworten der Probanden aufnehmen konnte.

Da der Name „Stefan“ in der späteren Ergebnisanalyse nicht gewertet wurde, gab es in jeder Testliste insgesamt 60 Wörter, die die Probanden richtig erkennen konnten.

Die Tests wurden alle in einem schallarmen Raum (2,9 m x 4 m) als Freifeldmessungen durchgeführt. In allen drei Experimenten wurde ein Lautsprecher (Event Electronics, 20/20bas V2) frontal vor dem Probanden in einem Abstand von 1,2 m aufgestellt.

In Experiment 3 wurde ein weiterer Lautsprecher (Event Electronics, 20/20bas V2) verwendet, der rechts oder links in einem Azimut-Winkel von + oder -90°, ebenfalls in einem Abstand von 1,2 m, vom Probanden entfernt stand. Die überlagerten Signale wurden kalibriert und immer bei einem Pegel von 70 dB SPL dargeboten. Dies wurde an der Position des Kopfes der Probanden gemessen.

### **8.6 Hör- und Messkonditionen**

Für Experiment 1 ergaben sich insgesamt 25 Messkonditionen (5  $\Delta F_0$  x 5 TMR). Jede Kondition wurde anhand einer Testliste geprüft. Die Messungen wurden bei unilateraler CI-Versorgung durchgeführt. Für den Fall, dass Probanden noch Restgehör auf der kontralateralen, nicht-implantierten Seite aufwiesen, wurde dieses Ohr mit einem Gehörschutzstöpsel (3M 1100) und einem zusätzlichen Gehörschutz-Kopfhörer (ENHA 3001) verschlossen.

Die Messungen dauerten ca. 1 ½ Stunden, zusätzlich individueller Pausen (Pyschny et al., 2009).

Die Stimuli von Experiment 2 bestanden aus 7 unterschiedlichen Maskierungskonditionen ( $\Delta F_0 = 40$  Hz, 80 Hz;  $\Delta F_n = 8\%$ , 16%;  $\Delta F_0 + \Delta F_n = 0$  Hz + 0%, 40 Hz + 8%, 80 Hz + 16%). Für jede Kondition wurden zwei Testlisten gemessen. Die Messungen wurden mit dem CI alleine, dem HG alleine und der Kombination CI + HG durchgeführt. Dies führte insgesamt zu 42 Messungen (7 Maskierungskonditionen x 2 Testlisten x 3 Hörhilfekonditionen). Die Nutzung der Hörhilfen und die Maskiererkonditionen wurden randomisiert.

Die Messungen wurden in zwei Sitzungen á ca. 2 Stunden aufgeteilt. Pausen wurden je nach Wunsch, Aufmerksamkeit und Konzentration der Probanden eingefügt (Pyschny et al., 2011).

## **Kapitel 8. Methode und Durchführung**

---

In Experiment 3 gab es 4 Maskierungskonditionen (Sprachmaskierer 100 Hz und 180 Hz, Rauschmaskierer 100 Hz und 180 Hz), die mit dem CI alleine, dem HG alleine und der Kombination CI + HG getestet wurden. Zusätzlich wurden die Maskierer aus drei verschiedenen Richtungen präsentiert (Azimut-Winkel: 0°, +90°, - 90°). Insgesamt ergaben sich somit 72 Messkonditionen (4 Maskierungskonditionen x 3 Maskiererorte x 3 Hörkonditionen x 2 Testlisten).

Die Messungen wurden in zwei Sitzungen aufgeteilt, die jeweils 2 - 3 Stunden dauerten. Zusätzlich wurden individuelle Pausen eingefügt.

### **8.7 Ethik**

Die folgenden Experimente wurden von der Ethik-Kommission der Universität zu Köln genehmigt und entsprechen der Deklaration von Helsinki.

## Kapitel 9

### Analyse und Statistik

In einer Kommunikationssituation mit konkurrierenden Sprechern, die sowohl durch IM als auch EM gekennzeichnet ist, können unterschiedliche Fehler auftreten. Die Unterschiede zwischen den verschiedenen Fehlertypen werden im Folgenden erläutert. Weiterhin werden die statistischen Untersuchungs

#### 9.1 Fehleranalyse

Fehler können entstehen, indem Hörer die Wörter des Ziel- und des Maskierersprechers miteinander verwechseln, so dass Wörter aus beiden Sätzen falsch zugeordnet werden. Beispielsweise könnte sich aus den Sätzen „Stefan kauft neun nasse Ringe“ und „Kerstin gibt zwei kleine Steine“ der Satz „Stefan kauft zwei kleine Ringe“ („zwei kleine“ anstatt „neun nasse“) ergeben. Dieser Fehler wird im weiteren Verlauf dieser Arbeit als A-Typ-Fehler bezeichnet.

Weiterhin können Fehler entstehen, die nicht auf Verwechslungen von Ziel- und Maskierersprecher beruhen. Hierbei handelt es sich um akustische Angleichungen an das Zielwort (z.B. „Rosen“ statt „Dosen“) oder das Auslassen von Wörtern (der Proband weiß das Wort nicht oder sagt nichts). Diese Fehler werden im weiteren Verlauf dieser Arbeit als B-Typ-Fehler bezeichnet.

Beide Fehlertypen sagen etwas über den Anteil von IM und EM aus. Nach Brungart (2001) sagen A-Typ-Fehler etwas über die Fähigkeit aus, Ziel- und Maskierersprecher voneinander trennen zu können. Es betrifft somit den Effekt der „stimulus uncertainty“ sowie der „target-masker similarity“. Dieser Fehlertyp wird dem IM zugerechnet. B-Typ-Fehler hingegen entstehen durch Signalreduktion, wenn die Energie des Maskierersignals stärker ist als die des Zielsignals und das Zielsignal deshalb nicht verstanden wird. B-Typ-Fehler können somit dem EM zugerechnet werden.

### 9.2 Auswertung

Die statistische Analyse wurde mithilfe der IBM Software SPSS (unterschiedliche Versionen seit 2007) durchgeführt. Im Ergebnisteil werden für jedes Experiment nur die statistisch signifikanten Effekte beschrieben und später diskutiert. Für weitere Informationen wird auf die Tabellen in Anhang C verwiesen.

#### 9.2.1 Experiment 1

Die Ergebnisse für das Sprachverstehen bei  $\Delta F_0$  und TMR der sechs CI-Probanden wurde mithilfe einer einfaktoriellen Varianzanalyse (ANOVA) ausgewertet. Bei einem signifikanten Ergebnis wurden Post-Hoc Tests durchgeführt. Die p-Werte der anschließenden Post-Hoc Tests wurden nach Bonferroni adjustiert und sind auf dem  $p = 0,05$ -Niveau signifikant.

Die Kombination der Faktoren TMR und  $\Delta F_0$  wurde nicht explizit gemessen. Um jedoch aufzeigen zu können, wie groß der Einfluss der einzelnen Faktoren TMR und  $\Delta F_0$  in der kombinierten Änderung war, wurde mithilfe einer Modellfunktion das Sprachverstehen berechnet (Pyschny et al., 2009). Dafür wurden zwei polynomische Funktionen 3. Grades auf Grundlage der gemessenen Ergebnisse ermittelt, einerseits in Abhängigkeit von  $\Delta F_0$  (bei TMR = 0 dB) und andererseits in Abhängigkeit von TMR (bei  $\Delta F_0 = 0$  Hz). Beide Funktionen wurden linear miteinander kombiniert:

$$\text{„SV [\%]} = 0,004 * \text{TMR}^3 - 0,3 * \text{TMR}^2 + 7,83 * \text{TMR} + \\ 0,00008 * \Delta f_0^3 - 0,008 * \Delta f_0^2 + 0,29 * \Delta f_0 + 29,43\text{“} \quad (\text{Pyschny et al., 2009: 674})$$

#### 9.2.2 Experiment 2

Da die Daten nicht normalverteilt waren, wurde eine Arkussinus-Transformation vorgenommen. Danach wurde eine einfaktorielle Varianzanalyse (ANOVA) mit Messwiederholung für die unterschiedlichen Hörhilfen (CI, HG und CI + HG) und für die unterschiedlichen Ziel-Maskierer-Konditionen ( $\Delta F_0$ ,  $\Delta F_n$  und  $\Delta F_0 + \Delta F_n$ ) durchgeführt.

Als Prüfgröße wurde Wilks Lambda verwendet. Die p-Werte der Post-Hoc Tests wurden nach Bonferroni adjustiert und sind auf dem  $p = 0,05$ -Niveau signifikant. Bedingt durch das Studiendesign konnte keine multifaktorielle ANOVA durchgeführt werden, da zwei Ziel-Maskierer-Konditionen ( $\Delta F_0 + \Delta F_n = 40$  Hz 16% und  $\Delta F_0 + \Delta F_n = 80$  Hz 8%) nicht generiert wurden.

### 9.2.3 Experiment 3

Da sich Experiment 3 mit den Effekten des räumlichen Hörens beschäftigt, soll zuerst aufgezeigt werden, wie aus bestimmten Hörkonditionen der Kopfschatten-, der Squelch- und der Summationseffekt berechnet wird. Diese Berechnungen folgen dem Beispiel von Schleich et. al. (2004). Abbildung 9.1 veranschaulicht die Hörsituationen, die für die Berechnung der drei Effekte erforderlich waren.

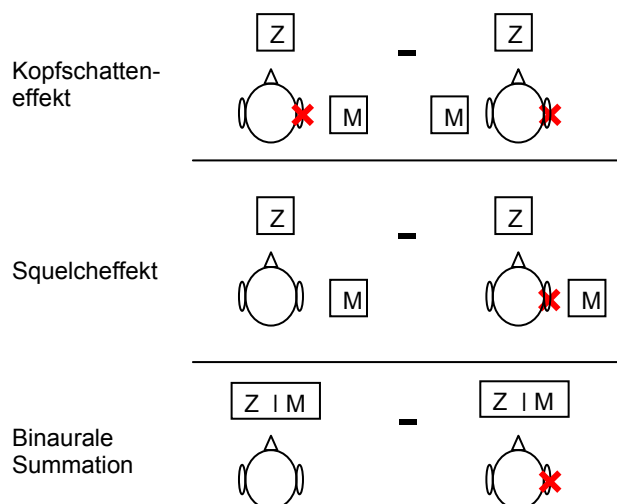


Abbildung 9.1: Hörkonditionen für die Berechnung der Effekte räumlichen Hörens Z = Ziel; M = Maskierer; rotes Kreuz = nicht genutztes Ohr

Der Kopfschatteneffekt wurde definiert als die Differenz der Zielverständlichkeit mit der Störquelle auf der Seite des unversorgten Ohres und der Zielverständlichkeit mit der Störquelle auf der Seite des versorgten Ohres. Die Störquelle befand sich in einem Azimut-Winkel von  $+ / - 90^\circ$  vom Hörer.

Häufig wird der Kopfschatteneffekt auf eine andere Art berechnet, die jedoch das Hören mit beiden Ohren voraussetzt. Sobald beide Ohren involviert sind, spielen immer auch binaurale Phänomene eine Rolle. In der hier verwendeten Messart kann das Ausmaß des Kopfschattens unabhängig von den weiteren Effekten Squelch und Summation betrachtet werden.

Der Squelcheffekt wurde berechnet als Differenz der Verständlichkeit bei bimodaler Versorgung und der Zielverständlichkeit für das Hören mit dem der Störquelle abgewandten Ohr. Die Störquelle befand sich in einem Azimut-Winkel von  $+ / - 90^\circ$  vom Hörer.

Die binaurale Summation ergab sich aus der Differenz des bimodalen und des unilateralen Hörens, wobei sowohl Nutz- als auch Störsignal aus frontaler Richtung (Azimut-Winkel =  $0^\circ$ ) präsentiert wurden.

Um aufzuzeigen, ob die drei Effekte Kopfschatten, Squelch und Summation zu signifikanten Unterschieden im Sprachverstehen führten, wurden für das Zielverstehen und die Fehlerverteilung T-Tests durchgeführt. Da es sich um Berechnungen von Differenzen handelte, die nahezu normal verteilt waren, wurde keine Standardisierung der Daten vorgenommen.

Mithilfe einer dreifaktoriellen ANOVA wurden die Unterschiede zwischen den Maskierertypen (Sprache vs. Rauschen), den Hörhilfen (CI vs. HG) und den beiden Grundfrequenzen (100 Hz vs. 180 Hz) für das Zielsatzverstehen getestet.

Da die Fehleranalyse nur für die Sprachmaskierer diskutiert wird, wurde die Fehlerverteilung mithilfe einer zweifaktoriellen ANOVA für die Hörhilfen und die Grundfrequenzunterschiede ausgewertet. Als Prüfgröße wurde Wilks Lambda verwendet. Ergaben sich innerhalb der ANOVA signifikante Effekte, wurden für diese Ergebnisse zusätzlich Post-Hoc Tests durchgeführt, deren p-Werte ebenfalls nach Bonferroni adjustiert und auf dem 0,05-Niveau signifikant sind.

## Kapitel 10

### Ergebnisse zu Experiment 1: Pegel- und Grundfrequenzunterschiede

Experiment 1 soll den Nachweis über die Fähigkeit unilateral versorgter CI-Träger erbringen, inwieweit Unterschiede zwischen einem männlichen Zielsprecher und einem simultan überlagerten Maskierersprecher wahrgenommen und genutzt werden können.

Das Signal des Maskierersprechers wurde aus dem Signal des Zielsprechers generiert. In der neutralen Kondition hatten Ziel- und Maskierersprecher die gleiche Stimme ( $\Delta F0 + TMR = 0 \text{ Hz} + 0 \text{ dB}$ ). Um Unterschiede zwischen den Stimmen einzufügen, wurde die Grundfrequenz des Maskierersprechers von 100 Hz bis 180 Hz in 20 Hz-Schritten verändert. Weiterhin wurden Pegelunterschiede eingefügt, die von 0 dB bis 20 dB in 5 dB-Schritten vorgenommen wurden. Der Maskierersprecher wurde im Vergleich zum Zielsprecher leiser.

Es wird erwartet, dass Pegelunterschiede besser genutzt werden können als Grundfrequenzunterschiede, um den Zielsprecher zu verstehen. Insgesamt sollten mehr Wörter verstanden werden und somit der Anteil des EM verringert werden. Die Wahrnehmung der Unterschiede zwischen den Sprechern sollte auch den Anteil der Wörter verringern, die zwischen Ziel- und Maskierersprecher verwechselt werden. Dadurch würde Informationale Masking reduziert werden.

#### 10.1 Zielverständlichkeit

In Abbildung 10.1 sind die gemittelten Ergebnisse der 6 CI-Probanden über den Grundfrequenzunterschieden für unterschiedliche TMR aufgetragen. Die Sprachverständlichkeit ist als %-korrekt angegeben. Die Fehlerbalken stehen für + / - eine Standardabweichung.



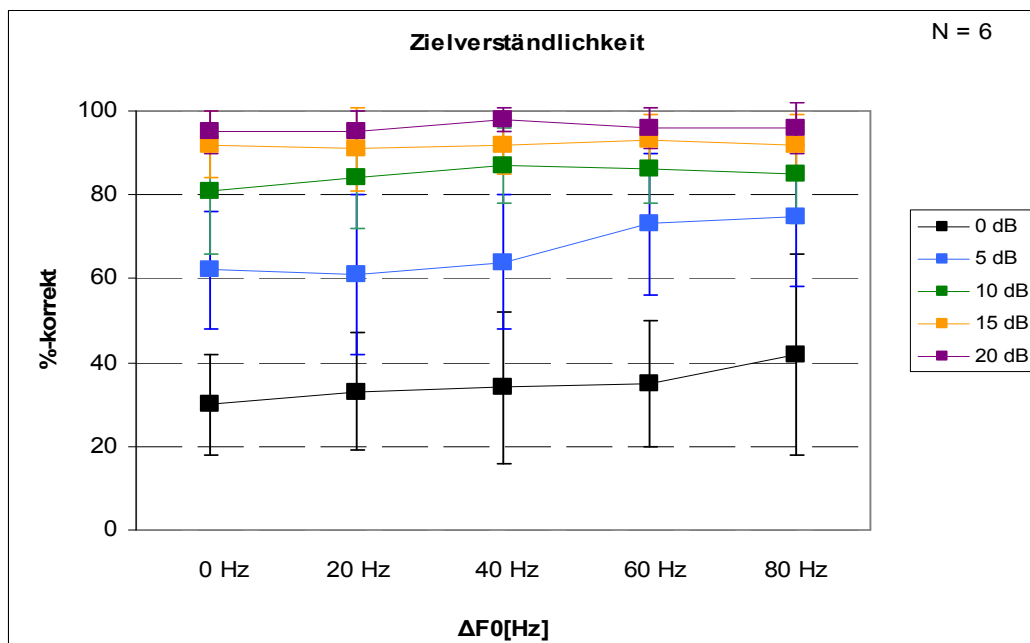


Abbildung 10.1: Zielverständlichkeit gemittelt über die 6 CI-Probanden (in Prozent) bei unterschiedlichen Pegel- und Grundfrequenzunterschieden zwischen Ziel- und Maskierersprecher.

In der schwierigsten Hörkondition, bei gleichem Ziel- und Maskierersprecher ( $\Delta F0 = 0 \text{ Hz} + \text{TMR} = 0 \text{ dB}$ ), wurden lediglich 30% des Zielsatzes verstanden. Die Variation in  $F0$  führte nur zu einer minimalen Verbesserung im Sprachverstehen. Ohne Pegelunterschiede zwischen Ziel- und Maskierersprecher (schwarze Linie:  $\text{TMR} = 0 \text{ dB}$ ) verbesserte sich das Sprachverstehen für den größten Unterschied in  $F0$  (80 Hz) um 13% (von 28% auf 41%). Bei einem geringen Pegelunterschied von 5 dB verbesserte sich die Sprachverständlichkeit mit Änderungen der Grundfrequenz ebenfalls um 13% (von 62% bei  $\Delta F0 = 0 \text{ Hz}$  auf 75% bei  $\Delta F0 = 80 \text{ Hz}$ ). Bei höheren TMR ab 10 dB hatte die Änderung der Grundfrequenz keinen Effekt mehr. Insgesamt belief sich der Nutzen durch die Grundfrequenzunterschiede gemittelt auf nur 0,15% / Hz.

Ein Unterschied im Pegel führte hingegen zu einem deutlichen Zugewinn im Sprachverstehen der Probanden. In der schwierigsten Hörkondition, ohne  $F0$ -Unterschied zwischen den Stimmen, führte die kleinste Veränderung im Pegel ( $\text{TMR} = 5 \text{ dB}$ ) zu einer Verbesserung um ca. 30% im Sprachverstehen. Das entspricht einem Zugewinn von 6% / dB. Ab einem Pegelunterschied von 15 dB kam es zu einem Deckeneffekt.

Eine einfaktorielle Varianzanalyse (ANOVA) wies nur einen signifikanten Effekt für TMR ( $p < 0,001$ ) auf. Weder eine Änderung in  $F0$  ( $p = 0,376$ ) noch die Kombination aus beiden Parametern  $\Delta F0 \times \text{TMR}$  ( $p = 0,995$ ) führte zu einem signifikant unterschiedlichen Sprachverstehen (s. Tabelle C 1.1 in Anhang C.1).

## Kapitel 10. Ergebnisse zu Experiment 1: Pegel- und Grundfrequenzunterschiede

Sowohl die Ergebnisse bei einem TMR von 0 dB als auch die Ergebnisse bei einem TMR von 5 dB waren für  $\Delta F_0 = 0$  Hz von den Ergebnisse aller übrigen TMR signifikant unterschiedlich ( $p < 0,001$ ).

### 10.2 Fehlerverteilung

Wie bereits in Kapitel 9 beschrieben, ergeben sich zwei unterschiedliche Fehlertypen, durch die eine präzisere Analyse hinsichtlich des Anteils von IM und EM durchgeführt werden kann.

#### 10.2.1 Anteil der A-Typ-Fehler

Abbildung 10.2 zeigt den Anteil der A-Typ-Fehler über  $\Delta F_0$  für die verschiedenen TMR. Dargestellt ist der Anteil der verwechselten Wörter in Prozent. Die Fehlerbalken stehen für  $\pm$  eine Standardabweichung.

Der Anteil der A-Typ-Fehler wird für TMR = 0 dB hervorgehoben dargestellt (schwarze Linie: TMR = 0 dB), um gezielt den Einfluss der Grundfrequenz erläutern zu können.

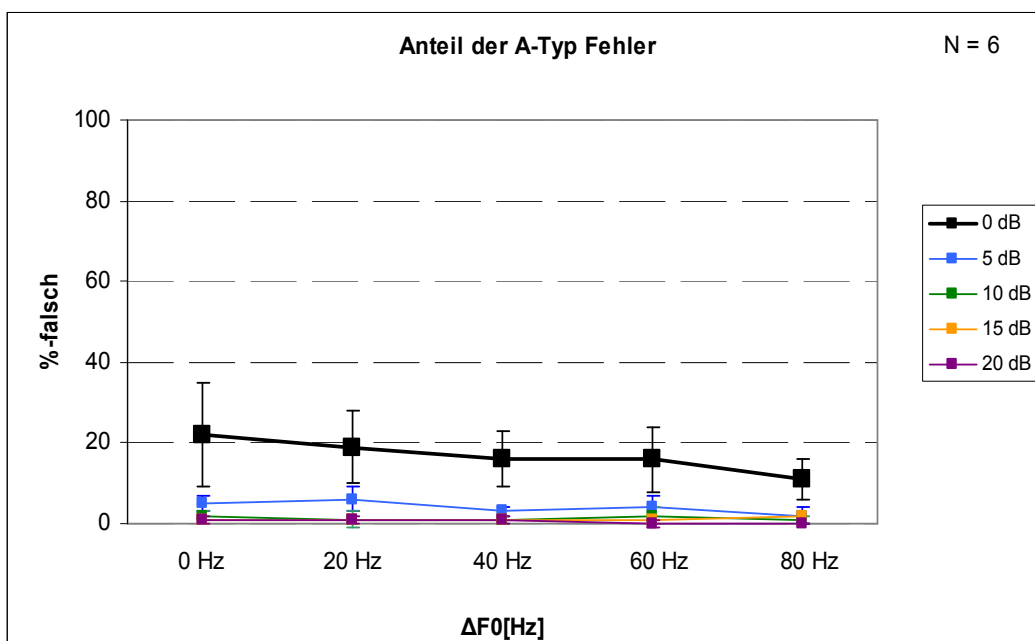


Abbildung 10.2: Anteil der A-Typ-Fehler gemittelt über die 6 CI-Probanden (in Prozent) bei unterschiedlichen Pegel- und Grundfrequenzunterschieden zwischen Ziel- und Maskierersprecher.

Bei identischem Ziel- und Maskierersprecher ( $\Delta F_0 = 0$  Hz + TMR = 0dB) wurden 22% der Wörter verwechselt. Mit steigender  $F_0$  des Maskierers und damit zunehmender Differenz in

## Kapitel 10. Ergebnisse zu Experiment 1: Pegel- und Grundfrequenzunterschiede

F0 sanken die A-Typ-Fehler auf 11% beim größten F0-Unterschied von 80 Hz. Bei den restlichen TMR blieben die A-Typ-Fehler konstant zwischen 0% und 5%.

Die einfaktorielle ANOVA zeigte für den A-Typ-Fehler nur einen signifikanten Effekt für TMR ( $p < 0,001$ ). Ein Unterschied in F0 ( $p = 0,073$ ) oder in der Kombination  $\Delta F0 \times TMR$  ( $p = 0,270$ ) führte zu keiner signifikanten Änderung der Verwechslungsfehler. Die Post-Hoc Tests zeigten bei gleicher Grundfrequenz von Ziel- und Maskierersprecher, dass ein TMR von 0 dB von allen restlichen TMRs unterschiedlich war. Ein TMR von 5 dB unterschied sich von einem TMR von 20 dB ( $p = 0,036$ ) (s. Tabelle C 1.2 in Anhang C.1).

### 10.2.2 Anteil der B-Typ-Fehler

Abbildung 10.3 zeigt den Anteil der B-Typ-Fehler aufgetragen über  $\Delta F0$  für die verschiedenen TMR. Dargestellt ist der Anteil der falschen Antworten in Prozent. Die Fehlerbalken stehen für  $\pm$  eine Standardabweichung. Auch hier wird das Hauptaugenmerk auf die Kondition TMR = 0 dB gelegt, die grafisch (schwarze Linie) hervorgehoben wurde.

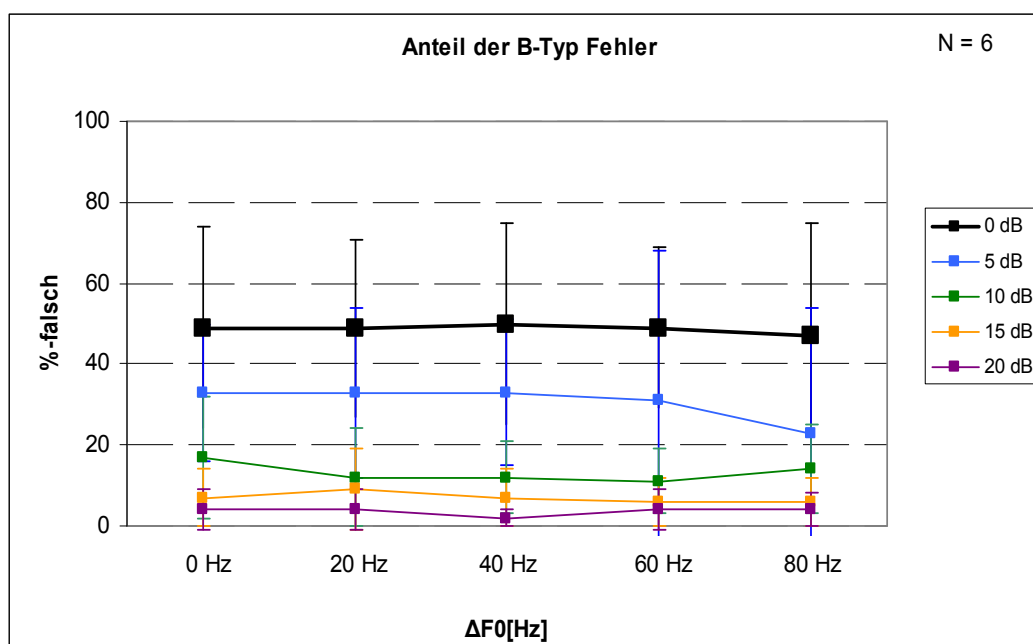


Abbildung 10.3: Anteil der B-Typ-Fehler gemittelt über die 6 CI-Probanden (in Prozent) bei unterschiedlichen Pegel- und Grundfrequenzunterschieden zwischen Ziel- und Maskierersprecher.

Insgesamt zeigte sich, dass mit steigendem TMR die B-Typ-Fehler bei allen getesteten Grundfrequenzunterschieden reduziert wurden.

Bei TMR = 0 dB blieb der Anteil der B-Typ-Fehler über alle Grundfrequenzunterschiede nahezu konstant, bei im Mittel 48%. In der Kondition TMR = 5 dB sanken die B-Typ-Fehler

## **Kapitel 10. Ergebnisse zu Experiment 1: Pegel- und Grundfrequenzunterschiede**

bis zu dem größten F0-Unterschied von 80 Hz, von 33% auf 23% ab. Für höhere TMR ergaben die Ergebnisse einen Bodeneffekt.

Die einfaktorielle ANOVA ergab auch für den B-Typ-Fehler nur einen signifikanten Unterschied für TMR ( $p < 0,001$ ). Ein Unterschied in F0 ( $p = 0,764$ ) oder in der Kombination  $\Delta F0 \times TMR$  ( $p = 0,998$ ) führte zu keiner signifikanten Änderung der falsch oder nicht verstandenen Wörter. Bei gleicher Grundfrequenz waren die kleinsten TMR von 0 dB und 5 dB jeweils von allen anderen TMRs signifikant unterschiedlich ( $p < 0,001$ ) (s. Tabelle C 1.3 in Anhang C.1).

### **10.3 Diskussion**

In diesem ersten Experiment wurde untersucht, inwieweit unilateral versorgte CI-Träger einen Zielsprecher verstehen, wenn dieser von einem simultan dargebotenen Maskierersprecher verdeckt wurde. Es wurde überprüft, ob Pegel- und / oder Grundfrequenzunterschiede zwischen den Sprechern dazu beitragen, IM in einer solchen Situation zu reduzieren.

Alleinige Unterschiede in der Grundfrequenz konnten von den CI-Trägern kaum genutzt werden, um den Zielsprecher besser zu verstehen und konkurrierende Sprecher voneinander zu trennen. Es zeigte sich nur eine Verbesserung bei dem größten F0-Unterschied, der in Kombination mit großen Pegelunterschieden jedoch nicht mehr zu beobachten war.

Meister et al. (2009) untersuchten bei CI-Trägern die Wahrnehmung prosodischer Merkmale, die auf der Wahrnehmung der Grundfrequenz basierten. Neben Experimenten zur Erkennung von Frage vs. Aussagesätzen und zur Satzbetonung wurde ein Experiment zur Erkennung des Sprechergeschlechts durchgeführt. Es handelte sich um sechs natürliche Sprecher (drei männlich, drei weiblich), die den Grundfrequenzbereich von einer tiefen Männerstimme (85 Hz) bis zu einer hohen Frauenstimme (270 Hz) abdeckten. Es zeigte sich, dass die CI-Träger bei der Erkennung eines männlichen Sprechers mit einer Grundfrequenz von 165 Hz Probleme hatten. Die Stimme wurde nicht einheitlich als männlich oder weiblich identifiziert.

Wahrscheinlich wurden die konkurrierenden Sprecher in dem vorliegenden Experiment in einem Bereich von 100 Hz bis 160 Hz als männlich und somit als zum gleichen Sprechergeschlecht zugehörig wahrgenommen. Die Grundfrequenzunterschiede reichten alleine nicht aus, um die Sprecher trennen zu können. In diesem Bereich bedarf es zusätzlich der Wahrnehmung höherer Frequenzen, um die Formantfrequenzen nutzen zu können. Diese Merkmale werden jedoch von CIs ebenfalls nur eingeschränkt übertragen und können nur schlecht von den CI-Trägern wahrgenommen werden.

## **Kapitel 10. Ergebnisse zu Experiment 1: Pegel- und Grundfrequenzunterschiede**

Der Unterschied von 100 Hz zu 180 Hz in dem vorliegenden Experiment entsprach, anlehnend an die Studie von Meister et al. (2009), einem Unterschied im Sprechergeschlecht. Wie aus dem Vortest ersichtlich (s. Kapitel 7), konnte das Sprechergeschlecht durch Änderungen in F0 erkannt und unterschieden werden. In der konkurrierenden Sprechersituation eigneten sich die Änderungen in F0 nur bedingt, um Sprecher voneinander zu trennen und IM zu reduzieren. Auch wenn das Ergebnis statistisch nicht signifikant war, ist es für die klinische Routine dennoch interessant. Es bestand die Tendenz einer verbesserten Sprachverständlichkeit mit zunehmenden F0-Unterschieden zwischen den Sprechern. Dies ist auch anhand abnehmender Verwechslungen (A-Typ-Fehler) ersichtlich. A-Typ-Fehler stehen in Zusammenhang mit IM (Brungart, 2001). Die Verringerung würde somit für eine Reduktion von IM sprechen und auf eine, wenn auch eingeschränkte, Fähigkeit zur Trennung von Sprechern bei großen F0-Unterschieden hinweisen.

Falsch oder nicht verstandene Wörter (B-Typ-Fehler) änderten sich hingegen durch Änderungen in F0 nicht. Da B-Typ-Fehler in Zusammenhang mit EM gesehen werden, bedeutet das, dass sich der Anteil von EM durch Unterschiede in F0 nicht veränderte. Als Erklärung für die Ergebnisse dient hier die eingeschränkte spektrale und temporale Verarbeitung im CI (Stickney et al., 2004). Kleine Unterschiede im Originalspektrum des Eingangssignals werden durch die Bandpassfilterung und die anschließende Abbildung auf eine begrenzte Anzahl von Elektrodenkontakten nicht aufgelöst (Rubinstein 2004).

Zeng (2002) konnte zeigen, dass es für CI-Träger schwierig war, Stimmen mit ähnlicher Grundfrequenz voneinander zu trennen. Weder durch zeitliche noch durch ortskodierte Mechanismen war die Abbildung der Grundfrequenz für CI-Träger stabil.

Hinzu kommt, dass auch spektrale Aspekte, wie die Formanten von Vokalen und den Transienten, durch die wenigen Kanäle nicht genügend abgebildet werden. CI-Träger haben somit nur bedingt Zugriff auf relevante Information von Stimmen, die benötigt werden, um Sprecher zu erkennen und voneinander zu unterscheiden (Vongphoe und Zeng, 2005).

Eine deutliche Verbesserung des Sprachverstehens wurde durch Unterschiede im Pegel zwischen den beiden Sprechern erreicht. Durch TMRs wurden sowohl IM als auch EM reduziert.

In der kombinierten Änderung von TMR und  $\Delta F0$  war die Annahme, dass sich das Sprachverstehen (SV) additiv verhält. Anhand der Modellfunktion (s. Kapitel 9.2.1) konnte gezeigt werden, dass Pegelunterschiede das Sprachverstehen mehr beeinflussten als Grundfrequenzunterschiede. Dies galt für höhere TMRs ab einem Wert von 5 dB (Pyschny et al., 2009).

Die Fehlerverteilung unterschied sich von der normalhörender Personen. Während bei NH die wesentlichen Fehler durch die Verwechslungen von Ziel- und Maskiererwörtern zustande

## **Kapitel 10. Ergebnisse zu Experiment 1: Pegel- und Grundfrequenzunterschiede**

kamen (Brungart, 2001), entstanden bei CI-Trägern die Fehler größtenteils dadurch, dass Wörter falsch oder nicht verstanden wurden. Anlehnend an diese Fehlerverteilung spielt bei den CI-Trägern EM offensichtlich eine größere Rolle als IM.

Insgesamt ist das Sprachverstehen von CI-Trägern in einer solchen Kommunikationssituation, verglichen mit dem Sprachverstehen von Normalhörenden, reduziert (Pyschny et al., 2009). NH sind in der Lage sowohl die kleinsten verfügbaren Pegel- als auch Grundfrequenzunterschiede zu nutzen, um EM und besonders IM zu reduzieren.

### **10.4 Zusammenfassung der Ergebnisse**

Insgesamt zeigte sich, dass CI-Träger Grundfrequenzunterschiede nur bedingt nutzen konnten, um konkurrierende Sprecher voneinander zu trennen und IM zu reduzieren. Nur durch den größten im Experiment verwendeten F0-Unterschied von 80 Hz verbesserte sich das Sprachverstehen aufgrund von sinkenden Verwechslungen. Die Anzahl der falsch oder nicht verstandenen Wörter wurden durch Grundfrequenzunterschiede nicht verringert. Unterschiede im Pegel führten hingegen zu einer deutlicheren Reduzierung des EM und des IM.

## Kapitel 11

### Ergebnisse zu Experiment 2: Grund- und Formantfrequenzunterschiede

Das zweite Experiment untersuchte die Fähigkeit bimodal versorgter Patienten, stimmliche Unterschiede zwischen Ziel- und Maskierersprechern wahrzunehmen. In der neutralen Kondition hatten Ziel- und Maskierersprecher die gleiche Stimme ( $\Delta F_0 + \Delta F_n = 0 \text{ Hz} + 0\%$ ). Unterschiede zum Zielsprecher wurden durch Änderungen der Maskiererstimme entweder durch eine unterschiedliche Grundfrequenz ( $\Delta F_0 = 40 \text{ Hz}$  oder  $80 \text{ Hz}$ ), durch Änderungen in den Formantfrequenzen ( $\Delta F_n = 8\%$  oder  $16\%$ ) oder durch die kombinierte Änderung von  $F_0$  und  $F_n$  ( $\Delta F_0 + \Delta F_n = 40 \text{ Hz} + 8\%$  oder  $80 \text{ Hz} + 16\%$ ) generiert. Es stellt sich die Frage, inwieweit stimmliche Parameter genutzt werden können, um eine Zielsprecher besser zu verstehen. Es besteht die Annahme, dass bei dem bimodal versorgten Probandenkollektiv das HG zu einer besseren Trennung der Sprecher und somit zu einer besseren Zielverständlichkeit führt. Nicht nur Energetic Masking, sondern auch Informationale Masking sollte mithilfe des HGs reduziert werden.

#### 11.1 Zielverständlichkeit

In Abbildung 11.1 ist das über die 12 bimodal versorgten CI-Probanden gemittelte Sprachverstehen für die unterschiedlichen Ziel-Maskierer Konditionen ( $\Delta F_0$ ,  $\Delta F_n$ ,  $\Delta F_0 + \Delta F_n$ ) für die Hörhilfen HG, CI und die Kombination CI + HG in Prozent aufgetragen. Die Fehlerbalken entsprechen +/- einer Standardabweichung. Als Inklusionskriterium galt ein gutes Sprachverstehen der Probanden. Dies war mit dem CI alleine besser als mit dem HG alleine. Deswegen wurde das Hauptaugenmerk auf dem zusätzlichen Beitrag des HGs zum CI gelegt. Es wurden vor allem die Hörkonditionen „unilateral CI“ und die „bimodale Kombination CI + HG“ miteinander verglichen. P-Werte, errechnet mithilfe der einfaktoriellen

## Kapitel 11. Ergebnisse zu Experiment 2: Grund- und Formantfrequenzunterschiede

ANOVA für die Unterschiede im Sprachverstehen zwischen den Hörhilfen, befinden sich in Tabelle C.2.1 im Anhang C.2. P-Werte für die Unterschiede in der Zielverständlichkeit, die sich aufgrund der Ziel-Maskierer-Konditionen ergaben, finden sich in Tabelle C 2.2.

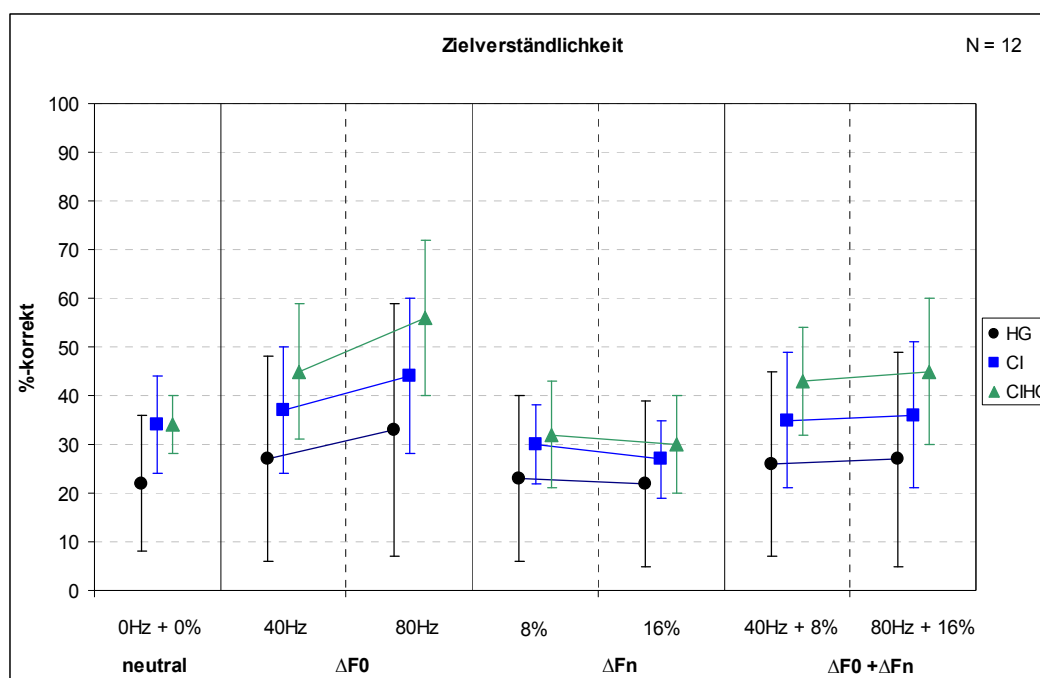


Abbildung 11.1: Durchschnittliches Sprachverstehen der 12 CI-Probanden (in Prozent) für das Hören mit dem HG, dem CI und der Kombination CI + HG in Prozent aufgetragen über die unterschiedlichen Ziel-Maskierer-Konditionen. Fehlerbalken stehen für +/- eine Standardabweichung.

Betrachtet man zunächst den Einfluss der Hörhilfen, zeigt sich insgesamt, dass die Probanden mit dem HG alleine eine Zielverständlichkeit zwischen 22% und 33% je nach Ziel-Maskierer-Kondition erreichten. Allerdings gab es besonders große Standardabweichungen, die zwischen 14% und 26% lagen, was für ein sehr heterogenes Bild bei der Nutzung der HGs spricht. Die hohe Standardabweichung kam u. a. dadurch zustande, dass 3 Probanden keinerlei Sprachverstehen mit dem HG in der konkurrierenden Sprechersituation zeigten.

Das Sprachverstehen in der unilateralen CI-Kondition und in der Kombination CI + HG war hingegen besser. Abhängig von der Ziel-Maskierer-Kondition verstanden die Probanden mit dem CI alleine zwischen 28% und 45%, in der bimodalen Kombination verstanden sie zwischen 30% und 55%. Gemittelt über alle Probanden und über alle Ziel-Maskierer-Konditionen gab es in der bimodalen Kondition jedoch nur einen Zugewinn von 6% gegenüber dem Verstehen mit dem CI alleine.

Der größte Zugewinn durch das HG zeigte sich in der Kondition  $\Delta F_0 = 80$  Hz. In der Kombination CI + HG verstanden die Probanden in dieser Kondition 12% mehr als mit dem CI alleine.



## **Kapitel 11. Ergebnisse zu Experiment 2: Grund- und Formantfrequenzunterschiede**

Die ANOVA zeigte signifikante Effekte beim Vergleich der Hörhilfen sowohl bei Änderungen von  $F_0$  um 40 Hz ( $p < 0,001$ ) und um 80 Hz ( $p < 0,001$ ) als auch in der kombinierten Änderung von  $F_0 + F_n$  bei 40 Hz + 8% ( $p = 0,003$ ) und 80 Hz + 16% ( $p = 0,002$ ). Anschließende Post-Hoc-Tests nach Bonferroni ergaben, dass in der bimodalen Kondition (CI + HG) das Sprachverstehen in den Konditionen  $\Delta F_0 = 40$  Hz ( $p = 0,009$ ),  $\Delta F_0 = 80$  Hz ( $p < 0,001$ ) und  $\Delta F_0 + \Delta F_n = 80$  Hz + 16% ( $p = 0,009$ ) signifikant besser war als in der unilateralen CI-Kondition.

Geht man nun zur Betrachtung des Einflusses der stimmlichen Variationen des Maskierersprechers über, zeigt sich insgesamt, dass das Sprachverstehen der Probanden bei Veränderungen von  $F_0$  und bei Veränderungen von  $F_0 + F_n$  gegenüber der neutralen Kondition ( $\Delta F_0 + \Delta F_n = 0$  Hz + 0%) anstieg. Die alleinige Änderung von  $F_n$  führte jedoch eher zu einer Verschlechterung des Sprachverstehens, besonders wenn mit dem CI alleine gehört wurde (s. Abbildung 11.1).

Die Ergebnisse der ANOVA (s. Tabelle C.2.2) zeigten, dass Änderungen in  $F_0$  zu einem besseren Sprachverstehen in der unilateralen CI-Kondition ( $p = 0,001$ ) und in der bimodalen Kondition CI + HG ( $p = 0,001$ ) führten. Bonferroni Post-Hoc-Tests zeigten auf, dass die Verbesserungen bei 80 Hz im Vergleich zur neutralen Kondition (CI:  $p = 0,036$ ; CI + HG:  $p = 0,003$ ) und im Vergleich  $\Delta F_0 = 40$  Hz zu  $\Delta F_0 = 80$  Hz (CI:  $p = 0,003$ ; CI + HG:  $p = 0,003$ ) auftraten.

Durch Verkürzung der Vokaltraktlänge und damit Erhöhung der  $F_n$  ergab sich ein signifikanter Effekt (ANOVA:  $p = 0,003$ ) für das Hören mit dem CI alleine. Nachfolgende Bonferroni-Tests zeigten, dass sich bei Änderungen von  $\Delta F_n$  um 16% das Sprachverstehen gegenüber der neutralen Kondition ( $\Delta F_0 + \Delta F_n = 0$  Hz + 0%) signifikant verschlechterte ( $p = 0,012$ ).

Bei Änderung beider Parameter ( $F_0 + F_n$ ) verbesserte sich das Sprachverstehen (ANOVA:  $p = 0,016$ ) für das Hören in der bimodalen Kondition. Der Bonferroni-Test zeigte, dass dies jedoch nur in der Kondition  $\Delta F_0 + \Delta F_n = 40$  Hz + 8% im Vergleich zur neutralen Kondition  $\Delta F_0 + \Delta F_n = 0$  Hz + 0% signifikant war.

### **11.2 Fehlerverteilung**

In Experiment 2 lässt sich durch die Verteilung der Fehlertypen feststellen, inwieweit das HG eher zu einer Reduzierung von Verwechslungen (A-Typ-Fehler) oder zu einer Reduzierung von falsch und/oder nicht verstandenen Wörtern (B-Typ-Fehler) führt.

**11.2.1 Anteil der A-Typ-Fehler**

Abbildung 11.2 zeigt den Anteil der A-Typ-Fehler in Prozent gemittelt über die 12 Probanden für das Hören mit dem HG, mit dem CI und mit der Kombination CI + HG für die unterschiedlichen Ziel-Maskierer-Konditionen  $\Delta F_0$ ,  $\Delta F_n$  und  $\Delta F_0 + \Delta F_n$ . Je höher der Wert, desto höher die Verwechslungen und desto schlechter das Sprachverstehen des Zielsatzes. Die Fehlerbalken stehen für +/- eine Standardabweichung. Die errechneten p-Werte sind in Tabelle C.2.1 für den Effekt der Hörhilfen und in Tabelle C.2.3 für den Einfluss der unterschiedlichen Ziel-Maskierer-Konditionen in Anhang C.2 aufgetragen.

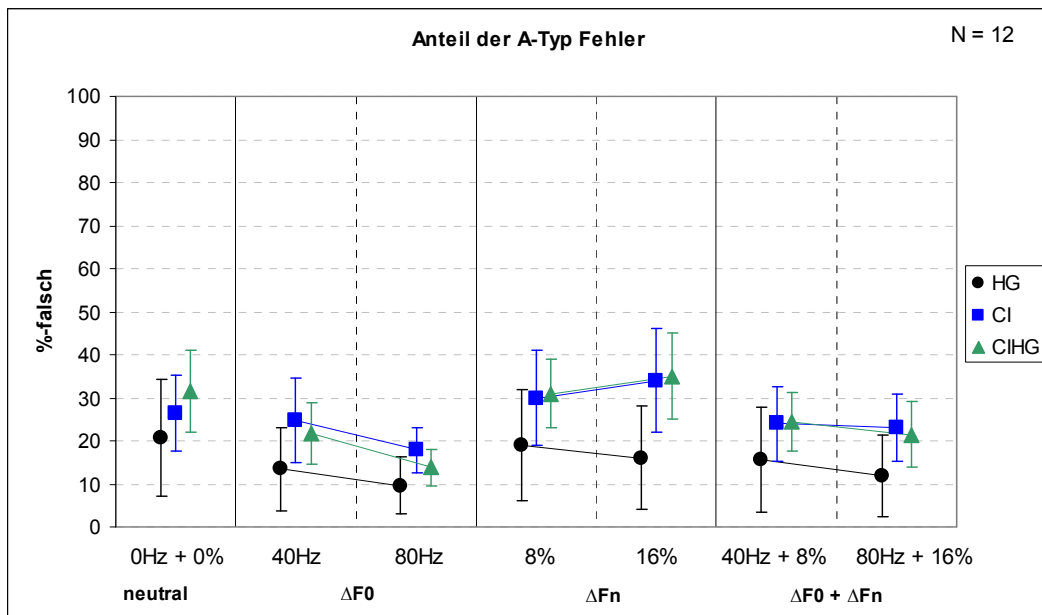


Abbildung 11.2: Durchschnittlicher Anteil der A-Typ-Fehler der 12 CI-Probanden (in Prozent) für die unterschiedlichen Ziel-Maskierer-Konditionen für das Hören mit dem HG, dem CI und der Kombination CI + HG. Fehlerbalken stehen für +/- eine Standardabweichung.

Allgemein veränderte sich der Anteil der A-Typ-Fehler für das Hören mit dem HG durch unterschiedliche Ziel-Maskierer-Konditionen in einem Bereich von 10% bis 20%. A-Typ-Fehler kamen häufiger vor, wenn mit dem CI alleine oder mit der Kombination CI + HG gehört wurde. In beiden Hörkonditionen zeigte sich eine Abhängigkeit des Anteils der A-Typ von den Ziel-Maskierer-Konditionen.

Bezüglich der Hörhilfen zeigte die ANOVA in nahezu allen Ziel-Maskierer-Konditionen eine signifikante Änderung (s. Tabelle C.2.1). Allerdings gab es nach den durchgeführten Bonferroni-Tests nur in der neutralen Ziel-Maskierer-Kondition ( $\Delta F_0 + \Delta F_n = 0 \text{ Hz} + 0\%$ ) einen signifikant höheren A-Typ-Fehler in der bimodalen Kombination gegenüber der unilateralen CI-Kondition ( $p = 0,018$ ).

Stimmliche Änderungen in der Maskiererstimme beeinflussten die Anzahl der A-Typ-Fehler unterschiedlich. Während eine Änderung in  $F_0$  bei allen drei Hörkonditionen zu weniger Verwechslungen führte, hatte eine Änderung von  $F_n$  eine Steigerung der A-Typ-Fehler in der bimodalen und der unilateral CI-Kondition zur Folge.

Die ANOVA zeigte, dass die Änderung in  $F_0$  für alle drei Hörkonditionen signifikant war (CI:  $p = 0,005$ ; HG:  $p = 0,004$ ; CI + HG:  $p < 0,001$ ). In der unilateralen CI-Kondition war der Anteil der Verwechslungen in der Kondition  $\Delta F_0 = 80$  Hz gegenüber der neutralen Kondition  $\Delta F_0 + \Delta F_n = 0$  Hz + 0% ( $p = 0,033$ ) und gegenüber der Kondition  $\Delta F_0 = 40$  Hz ( $p = 0,018$ ) signifikant reduziert. In der unilateralen HG Kondition war der Anteil der A-Typ-Fehler in der Kondition  $\Delta F_0 = 40$  Hz ( $p = 0,018$ ) und in der Kondition  $\Delta F_0 = 80$  Hz ( $p = 0,006$ ) gegenüber der neutralen Kondition signifikant unterschiedlich.

In der bimodalen Kondition führte eine Änderung um 40 Hz ( $p = 0,009$ ) und eine Änderung um 80 Hz ( $p < 0,001$ ) gegenüber der neutralen Kondition zu weniger Verwechslungen. Eine Änderung um 80 Hz führte im Vergleich zu einer Änderung von 40 Hz ebenfalls zu signifikant weniger Verwechslungen ( $p < 0,001$ ).

Bei der Durchführung der ANOVA für Änderungen in  $F_n$  gab es beim Hören mit dem CI alleine einen signifikanten Effekt ( $p = 0,028$ ), der nach Bonferroni-Korrektur allerdings nicht bestehen blieb.

Bei einer Änderung beider Parameter ( $\Delta F_0 + \Delta F_n$ ) wies die ANOVA in der unilateralen HG Kondition einen signifikanten Effekt ( $p = 0,008$ ) auf. In der Kondition  $\Delta F_0 + \Delta F_n = 80$  Hz + 16% ( $p = 0,015$ ) wurden gegenüber der neutralen Kondition weniger Wörter verwechselt. Auch in der bimodalen Hörkondition CI + HG wurden weniger Ziel- und Maskiererwörter verwechselt (ANOVA:  $p = 0,029$ ), allerdings war dieser Effekt nach Bonferroni-Korrektur nicht signifikant.

### **11.2.2 Anteil der B-Typ-Fehler**

Abbildung 11.3 zeigt den Anteil der B-Typ-Fehler in Prozent gemittelt über die 12 Probanden für alle drei Hörkonditionen für die unterschiedlichen Ziel-Maskierer-Konditionen  $\Delta F_0$ ,  $\Delta F_n$  und  $\Delta F_0 + \Delta F_n$ . Die Fehlerbalken stehen für +/- eine Standardabweichung. Die errechneten p-Werte für den Effekt der Hörhilfen sind in Tabelle C.2.1 dargestellt. Es gab keinen signifikanten Einfluss der unterschiedlichen Ziel-Maskierer-Konditionen.

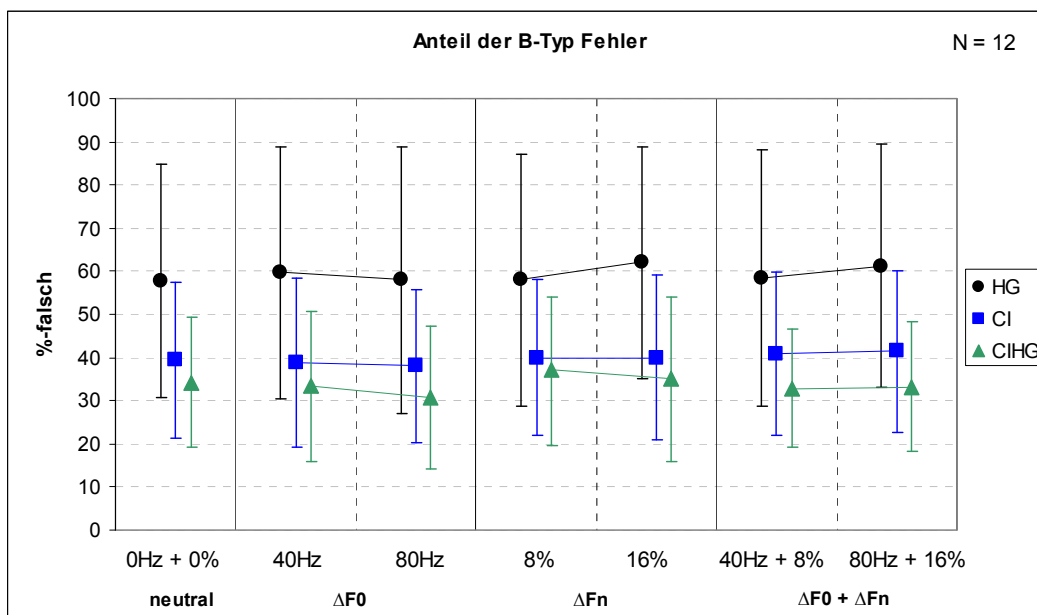


Abbildung 11.3: Durchschnittlicher Anteil der B-Typ-Fehler der 12 CI-Probanden (in Prozent) für die unterschiedlichen Ziel-Maskierer-Konditionen für das Hören mit dem HG, dem CI und der Kombination CI + HG. Fehlerbalken stehen für +/- eine Standardabweichung.

Insgesamt zeigte sich eine hohe Anzahl von B-Typ-Fehlern für das Hören mit dem HG alleine. Über alle Ziel-Maskierer-Konditionen lag der Fehleranteil durchweg bei 60%. Auch für die anderen beiden getesteten Hörsituationen (unilateral CI und bimodal CI + HG) blieb die Anzahl der B-Typ-Fehler über alle Ziel-Maskierer-Konditionen hinweg eher konstant. In der bimodalen Kondition (CI + HG) wurden im Vergleich zur unilateralen CI-Kondition allerdings durchweg weniger Wörter falsch oder nicht verstanden.

Die ANOVA zeigte für die Hörhilfen signifikante Effekte in allen Ziel-Maskierer-Konditionen. Der Bonferroni-Test zeigte, dass sich signifikant weniger B-Typ-Fehler in der bimodalen Kondition im Vergleich zur unilateralen CI-Kondition ergaben, bei einer Änderung von F<sub>0</sub> um 80 Hz ( $p = 0,036$ ) (s. Tabelle C.2.1).

Mittels der ANOVA konnte kein signifikanter Effekt der Ziel-Maskierer-Konditionen auf den B-Typ-Fehler festgestellt werden.

Zusammenfassend zeigten sich für die Hörkonditionen Unterschiede zwischen der unilateralen CI und der bimodalen Kondition bei F<sub>0</sub>-Unterschieden und großen Unterschieden in der Kombination F<sub>0</sub> + F<sub>n</sub>. Die Anzahl der A-Typ-Fehler unterschied sich nicht. Hingegen verringerte sich die Anzahl der B-Typ-Fehler in der Hörkondition CI + HG verglichen mit dem CI alleine, wenn eine große F<sub>0</sub>-Änderung vorlag.

Bezogen auf die stimmlichen Manipulationen der Maskiererstimme zeigte sich, dass eine Änderung in F<sub>0</sub> zu einem besseren Sprachverstehen und zu weniger Verwechslungen in den

Hörkonditionen CI alleine und CI + HG führte. Veränderung durch  $F_n$  führten eher zu einer Reduzierung des Sprachverstehens, besonders in der Kondition CI alleine.

### **11.3 Diskussion**

Im zweiten Experiment wurde der Einfluss der bimodalen Versorgung auf das Sprachverstehen in einer Kommunikationssituation mit zwei konkurrierenden Sprechern untersucht. Es wurde versucht, den Einfluss künstlich veränderter Grund- und Formantfrequenzen auf IM zu bestimmen.

Im Vergleich zum Verstehen mit dem CI alleine zeigte sich nur ein kleiner bimodaler Vorteil. Dieser kam vor allem bei großen Unterschieden in  $F_0$  und in  $F_0 + F_n$  zum Tragen. Kong et al. (2005) beobachteten einen Zugewinn durch das HG in einer Kommunikationssituation mit simultanen Sprechern. Die Studie belegte einen additiven Effekt von CI und HG speziell dann, wenn eine weibliche Stimme den männlichen Zielsprecher überlagerte, der Unterschied sowohl in  $F_0$  als auch in  $F_n$  also besonders groß war. Ihre Annahme war, dass der Zugewinn durch das HG aus der besseren Trennung von Ziel- und Maskierersprecher resultierte. Allerdings wurde keine explizite Fehleranalyse durchgeführt. Es ist aber anzunehmen, dass der Anteil der Verwechslung in ihrer Studie eher gering war, da immer derselbe Maskierersatz genutzt wurde und dieser den Hörern somit bekannt war. Im Gegensatz zu der Annahme von Kong et al. (2005) basierte der bimodale Zugewinn in der vorliegenden Studie nicht auf der Tatsache, dass Ziel- und Maskierersprecher besser getrennt werden konnten und dadurch die Verwechslungsfehler (A-Typ) vermindert wurden, sondern auf der Verringerung der falsch und nicht verstandenen Wörter (B-Typ-Fehler). Dies spricht eher für eine Reduzierung des EM als für eine Reduzierung des IM (Brungart, 2001). Würde das Hören in der bimodalen Kombination im Vergleich zum Hören mit dem CI alleine zu einer Reduzierung von IM führen, hätten die Verwechslungen zurückgehen müssen. Dies hätte gezeigt, dass das HG den Patienten hilft, Sprecher besser zu trennen (Pyschny et al., 2011). Durch die Reduzierung der B-Typ-Fehler erlangten die Probanden ein besseres allgemeines Sprachverstehen, was das Verstehen von sowohl Ziel- als auch Maskierersprecher beinhaltete.

Vor allem in den Konditionen mit Änderungen in  $F_0$  und Änderungen in  $F_0 + F_n$  zeigte sich der von Ching et al. (2007) beschriebene Komplementäreffekt des HGs zum CI. Die tieffrequente Information, z.B. der Grundfrequenz, übertragen durch das HG, vervollständigte die höheren Frequenzen, übertragen durch das CI (Cullington und Zeng, 2011), was zu einem besseren Sprachverstehen führte.

Mok et al. (2006) stellten in ihrer Studie fest, dass sich der größte Zugewinn in der bimodalen Kondition gegenüber der unilateralen CI-Kondition bei der Wahrnehmung von

## Kapitel 11. Ergebnisse zu Experiment 2: Grund- und Formantfrequenzunterschiede

bestimmten Phonemgruppen ergab. Laute, die ihre wichtigsten Merkmale im tieffrequenten Bereich haben, wie z.B. Diphthonge und Nasale, wurden mit CI + HG besser identifiziert als z.B. Frikative oder Plosive. Die Probanden waren in der Lage, die unterschiedliche Information beider Hörhilfen zu kombinieren. Anlehnend an diese Ergebnisse würde eine Reduzierung der B-Typ-Fehler in dieser Untersuchung eher auf eine Art phonemische Summation zurückgehen und weniger auf eine Reduktion des EM (Pyschny et al., 2011).

In der bimodalen Kondition kann das Phänomen des „Lückenhörens“ eine Erklärung für das bessere Sprachverstehen durch Änderungen in F0 sein (Cooke, 2005). Das „Lückenhören“ bezieht sich auf die Fähigkeit, weniger stark maskierte Bereiche bei konkurrierenden Signalen zu nutzen und dadurch Zielsignale besser zu verstehen. Eine Erklärung dafür lieferten Li und Loizou (2008). Sie präsentierten Normalhörenden Sprache, die vorher mit einem CI-Simulationsprogramm bearbeitet wurden. Die Sprachsignale wurden den Probanden entweder in dieser künstlichen Version alleine oder mit zusätzlichen tiefpassgefilterten Signalen vorgespielt, um somit eine Kombination aus akustischem und elektrischem Hören zu generieren. Ein Vorteil der akustisch-elektrischen Variante gegenüber der elektrischen Variante lag dann vor, wenn das Signal mit einem Sprecher überlagert wurde. Bei Überlagerung mit einem Rauschmaskierer zeigte sich diese Verbesserung nicht. Sie konnten einen sog. SNR-Vorteil im tieffrequenten Bereich aufzeigen. Dabei wurden die tieffrequenten Bereiche des Zielsprechers durch den simultanen Sprecher weniger maskiert als die hochfrequenten Bereiche. Dies traf z.B. zu, wenn in der Zielsprache stimmhafte Anteile enthalten waren wie z.B. Vokale und der Maskierersprecher dort weniger Energie aufwies. Dieser Effekt zeigte sich vor allem, wenn Ziel- und Maskierersprecher große Unterschiede in der Grundfrequenz aufwiesen, und könnte eine Erklärung für den positiven Effekt durch das HG bei  $\Delta F0 = 80$  Hz in der vorliegenden Untersuchung sein.

Hinsichtlich der stimmlichen Manipulationen führten besonders große Änderungen in der Grundfrequenz ( $\Delta F0 = 80$  Hz) zu einem besseren Sprachverstehen sowohl mit dem CI alleine als auch in der Kombination CI + HG. Anhand der sinkenden Verwechslungsfehler zeigte sich, dass es hier zu einer Verringerung von IM kam. Die Verringerung der Verwechslungen erhöhte sich jedoch in der bimodalen Kondition nicht im Vergleich zum Hören mit dem CI alleine. Das bedeutet, dass das Hörgerät keine zusätzliche Verringerung des A-Typ-Fehlers erbrachte (Pyschny et al., 2011).

Dies bestärkt die Ergebnisse von Experiment 1, in der die Tendenz zum besseren Sprachverstehen mit dem CI alleine durch große Änderungen in F0 ebenfalls durch reduzierte Verwechslungen hervorgerufen wurde.

Bei Änderung der Formantfrequenzen zeigte sich, dass die Zielverständlichkeit aufgrund von ansteigenden Verwechslungen zurückging. Dies scheint auf den ersten Blick verwunderlich, da durch die ansteigenden Formantfrequenzen des Maskierersprechers die

## **Kapitel 11. Ergebnisse zu Experiment 2: Grund- und Formantfrequenzunterschiede**

Unterschiede zwischen Ziel und Maskierer größer werden und somit eine bessere Trennung möglich sein sollte (Darwin et al., 2003). Die Ergebnisse könnten anlehnd an eine Studie von Cullington und Zeng (2008) erklärt werden. Diese untersuchten das Sprachverstehen von CI-Trägern bei konkurrierenden Sprechern und nutzten u. a. Kinderstimmen als Maskierer. Im Gegensatz zu den übrigen männlichen und weiblichen Maskierern entstand bei den Kinderstimmen mehr IM, obwohl deren Stimmen eine höhere Grundfrequenz hatten und somit einen größeren Unterschied zu dem männlichen Zielsprecher aufwiesen. Untersuchungen bzgl. der Sprechgeschwindigkeit oder Variationen in der Grundfrequenz der Kinderstimmen führten zu keiner Erklärung. Eine alleinige Veränderung der Formantfrequenzen lässt eine Stimme subjektiv kindlicher klingen. Verbunden mit den hier aufgeführten Ergebnissen könnte eine Erklärung sein, dass eben diese Veränderungen der Formantfrequenzen mehr IM hervorriefen und dadurch Sprecher nicht besser getrennt werden konnten (Pyschny et al., 2011).

Auch wenn der Anstieg der Verwechslungen bei  $\Delta F_n$  für das Hören mit dem CI allein nicht signifikant war, sind diese Beobachtungen von großer alltäglicher und klinischer Relevanz. Stimmen gleichen Geschlechts haben ähnliche Grundfrequenzen und unterscheiden sich eher durch die Formung des Vokaltrakts und somit veränderter Formantfrequenzen. Die Ergebnisse von Experiment 2 veranschaulichen, dass die eingeschränkte spektrale Auflösung der CIs dazu führt, dass Stimmen gleichen Geschlechts in einer konkurrierenden Sprechersituation nicht voneinander getrennt werden können. Dies entspricht den Ergebnissen des Vortests (Kapitel 7), bei dem die Wahrnehmung spektraler Unterschiede (z.B.  $F_n$ ) mit dem CI kaum gelang.

In der bimodalen Kondition verschlechterte sich das Sprachverstehen durch Änderungen in  $F_n$  nicht. Auch hier könnte die Erklärung eine Art phonemische Summation sein (Mok et al., 2006).

### **11.4 Zusammenfassung der Ergebnisse**

Das Hören in der bimodalen Kondition verbesserte das Sprachverstehen im Vergleich zum Hören mit dem CI alleine nur wenig. Die Annahme, dass IM durch die Hinzunahme des HGs reduziert wird und es zu einer besseren Trennung von Ziel- und Maskierersprecher kommt, bestätigte sich nicht. Verwechslungen zwischen den Sprechern wurden bei großen F0-Unterschieden zwar minimiert, dies galt jedoch gleichermaßen für die unilaterale CI-Kondition als auch für die bimodale Kondition. Das HG trug somit nicht gesondert zu einer besseren Differenzierung bei. Jedoch wurden aufgrund komplementärer Sprachinformation, übertragen durch das HG, weniger Wörter falsch oder gar nicht verstanden. Dies entspricht eher einer Verminderung des EM (Brungart, 2001).

## **Kapitel 11. Ergebnisse zu Experiment 2: Grund- und Formantfrequenzunterschiede**

Hinsichtlich der stimmlichen Parameter führten große Unterschiede in  $F_0$  zu einem besseren Sprachverstehen. Dies war bei Änderungen von  $F_n$  nicht der Fall.



## Kapitel 12

### Ergebnisse zu Experiment 3: Effekte räumlichen Hörens

In Experiment 3 wurde die Fähigkeit bimodal versorgter CI-Träger untersucht, Effekte des räumlichen Hörens in einer konkurrierenden Sprechersituation zu nutzen. Hierfür wurden zwei unterschiedliche Maskierer (Sprache und Rauschen) verwendet. Der Sprachmaskierer hatte entweder eine Grundfrequenz von 100 Hz oder von 180 Hz. Der modulierte Rauschmaskierer wurde auf der Grundlage der Sprachmaskierer, ebenfalls auf der Basis von 100 Hz und 180 Hz generiert.

Der Zielsprecher wurde immer aus einem 0° Azimut-Winkel (vor dem Hörer) präsentiert. Die Störsignale wurden entweder ebenfalls von vorne oder von rechts oder links aus einem Azimut-Winkel von + / - 90° dargeboten.

Es wird erwartet, dass durch die räumliche Trennung IM und EM reduziert werden. Hierbei rückt die Bedeutung der unterschiedlichen Grundfrequenz des Ziel- und des Störsignals in den Hintergrund. Weiterhin sollte binaurales Hören möglich sein, was zu einem besseren Sprachverstehen führt.

#### 12.1. Zielverständlichkeit

Abbildung 12.1 und 12.2 zeigen die Ergebnisse für die drei Effekte Kopfschatten, Squelch und Summation. Dargestellt ist die Differenz der Zielverständlichkeit in Prozent für die Verwendung der Rausch- (R) und Sprachmaskierer (S) mit unterschiedlicher F0 gemittelt über alle 10 Probanden. Die Fehlerbalken indizieren das 95%ige Konfidenzintervall. Positive Werte stehen für Verbesserungen in der Zielsprecherverständlichkeit, negative Werte für Verschlechterungen. Ein Stern über den Säulen zeigt an, ob der jeweilige Effekt signifikant ist. P-Werte können in Anhang C.3 nachgelesen werden, sind jedoch nicht explizit im Text genannt (Tabelle C 3.1). Signifikante Unterschiede zwischen den Maskierertypen (S und R)

## Kapitel 12. Ergebnisse zu Experiment 3: Effekte räumlichen Hörens

und F0, die sich durch die 3-faktorielle ANOVA errechneten, sind unterhalb der Säulen durch einen Stern verbunden. Unterschiede zwischen den Hörhilfen konnten nicht innerhalb einer Abbildung dargestellt werden. Soweit Unterschiede vorhanden sind, werden diese im Text beschrieben. Die Werte der ANOVA finden sich zusätzlich in den Tabellen C.3.2 – C.3.4 in Anhang C.3.

In Abbildung 12.1 ist der Kopfschatteneffekt für das CI berechnet. Der Squelcheffekt zeigt die Veränderung im Sprachverstehen für die Situation auf, in der das Störsignal von der Seite des CIs dargeboten wird und dies zum HG-Ohr hinzugeschaltet wird. Bei der Summation wird die Veränderung aufgezeigt, die die Hinzunahme des CIs zum HG bewirkt. Abbildung 12.2 ist dementsprechend für die Effekte des HG bzw. dessen Hinzunahme aufgebaut.

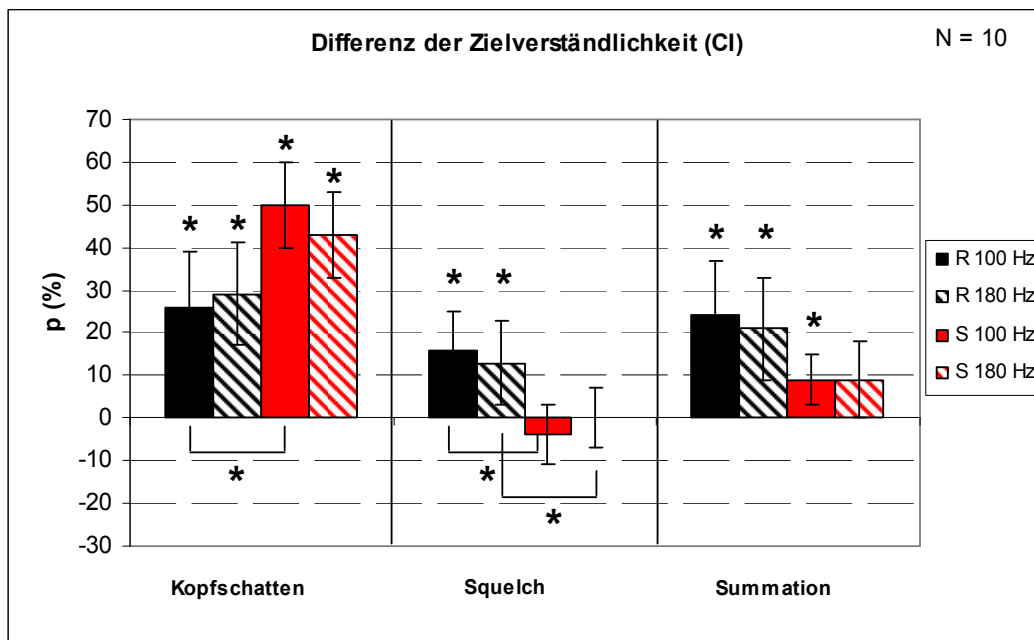


Abbildung 12.1: Differenz der Zielverständlichkeit (für CI) in Prozent aufgetragen für Rausch- und Sprachmaskierer für unterschiedliche Grundfrequenzen (F0 = 100 Hz; 180 Hz). Berechnungen: Kopfschatten für CI; Squelch: Störsignal auf Seite des CIs; Summation: Hinzunahme des CIs.

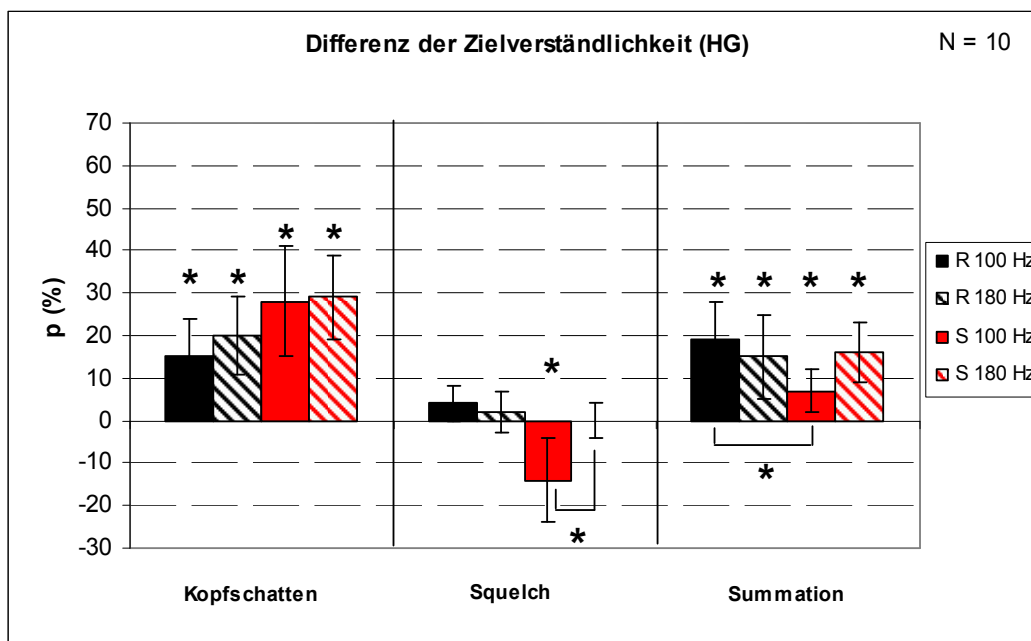


Abbildung 12.2: Differenz der Zielverständlichkeit (für HG) in Prozent aufgetragen für Rausch- und Sprachmaskierer für unterschiedliche Grundfrequenzen ( $F_0 = 100$  Hz;  $180$  Hz). Berechnungen: Kopfschatten für HG; Squelch: Störsignal auf der Seite des HGs; Summation: Hinzunahme des HGs.

**Kopfschatten**

Aufgrund des Kopfschattens verbesserte sich die Zielverständlichkeit für beide Hörhilfen sowohl bei der Verwendung der Rausch- als auch der Sprachmaskierer signifikant. Bei Verwendung der Rauschmaskierer verbesserte sich das Sprachverstehen für das Hören mit dem CI (s. Abbildung 12.1) um 26% ( $F_0 = 100$  Hz) bzw. um 29% ( $F_0 = 180$  Hz). Bei Verwendung der Sprachmaskierer war der Kopfschatteneffekt im Vergleich zu den Rauschmaskierern größer. Die Verbesserung lag bei 50% ( $F_0 = 100$  Hz) bzw. bei 43% ( $F_0 = 180$  Hz). Für das Hören mit dem HG (s. Abbildung 12.2) verbesserte sich das Sprachverstehen bei Verwendung der Rauschmaskierer um 15% ( $F_0 = 100$  Hz) bzw. um 20% ( $F_0 = 180$  Hz). Der Kopfschatteneffekt bei Verwendung der Sprachmaskierer war ebenfalls größer und lag bei 28% ( $F_0 = 100$  Hz) bzw. bei 29% ( $F_0 = 180$  Hz).

Bezüglich der Maskierertypen,  $F_0$  und der Hörhilfen wiesen die Ergebnisse der ANOVA (s. Tabelle C.3.2) auf einen signifikanten Effekt bei den Maskierertypen ( $p = 0.008$ ) hin. Insgesamt zeigte sich auch eine signifikante Interaktion zwischen Maskierertyp und  $F_0$  ( $p = 0.005$ ). Post-Hoc-Tests ergaben, dass der Kopfschatteneffekt für das Hören mit dem CI ( $p = 0.044$ ) bei Verwendung der Sprachmaskierer in der Kondition  $F_0 = 100$  Hz größer ausfiel als bei Verwendung der Rauschmaskierer ( $F_0 = 100$  Hz) (s. Abbildung 12.1).

### Squelch

Die Berechnung des Squelcheffekts ergab bei Verwendung der Rauschmaskierer, die sich ipsilateral auf der Seite des CIs befanden, eine signifikante Verbesserung von je ca. 15%. Hingegen traten bei Verwendung der Sprachmaskierer keine signifikanten Änderungen auf, wenn der Sprachmaskierer von der Seite des CIs präsentiert wurde (s. Abbildung 12.1).

Befand sich die Störquelle auf der Seite des HGs, ergaben sich keine nennenswerten Änderung in der Sprachverständlichkeit bei der Verwendung von Rauschmaskierern. Bei der Verwendung des Sprachmaskierers mit  $F_0 = 100$  Hz sank die Sprachverständlichkeit sogar signifikant um ca. 14% (s. Abbildung 12.2).

Die ANOVA (s. Tabelle C.3.3) zeigte einen signifikanten Unterschied zwischen den Maskierertypen (Rauschen vs. Sprache;  $p < 0.001$ ) und in der Interaktion von Maskierertyp und  $F_0$  ( $p = 0.002$ ). War die Störquelle auf der Seite des CIs angeordnet (s. Abbildung 12.1), zeigten Post-Hoc-Tests nach Bonferroni, dass der Squelcheffekt für beide Rauschmaskierer ( $F_0 = 100$  Hz,  $p = 0.016$ ;  $F_0 = 180$  Hz,  $p = 0.02$ ) größer ausfiel als für die Sprachmaskierer mit der jeweiligen Grundfrequenz. Befand sich die Störquelle auf der Seite des HGs, war das Sprachverstehen bei Verwendung der Sprachmaskierer in der Kondition  $F_0 = 100$  Hz im Vergleich zu der Kondition  $F_0 = 180$  Hz schlechter ( $p = 0.04$ ) (s. Abbildung 12.2). Für  $F_0$  wurde insgesamt in der ANOVA jedoch kein signifikanter Effekt festgestellt.

### Summation

Die Berechnung des Summationseffekts ergab ein signifikant verbessertes Sprachverstehen durch Hinzunahme des CIs im Vergleich zum Hören mit dem HG alleine (s. Abbildung 12.1) bei der Verwendung von Rauschmaskierern. Der Zugewinn belief sich auf einen Wert von 24% in der Kondition  $F_0 = 100$  Hz und von 21% in der Kondition  $F_0 = 180$  Hz. Bei der Verwendung der Sprachmaskierer verbesserte sich das Sprachverstehen um jeweils 9%. Allerdings war diese Verbesserung nur in der Kondition  $F_0 = 100$  Hz signifikant (s. Tabelle C.3.1 in Anhang C.3).

Durch die Hinzunahme des HGs verbesserte sich das Sprachverstehen bei Verwendung der Rauschmaskierer signifikant im Vergleich zum Hören mit dem CI alleine (s. Abbildung 12.2) um 19% in der Kondition  $F_0 = 100$  Hz und um 15% in der Kondition  $F_0 = 180$  Hz. Bei Verwendung der Sprachmaskierer ergab sich ebenfalls eine signifikante Verbesserung sowohl in der Kondition  $F_0 = 100$  Hz (7%) als auch in der Kondition  $F_0 = 180$  Hz (16%) (s. Tabelle C.3.1).

Die ANOVA (s. Tabelle C.3.4) zeigte einen signifikanten Unterschied zwischen den Maskierertypen auf ( $p = 0.001$ ). Die nachfolgend durchgeführten Post-Hoc Tests belegten, dass durch die Hinzunahme des HGs das Sprachverstehen im Vergleich zum Hören mit dem CI alleine (s. Abbildung 12.2) bei Verwendung des Rauschmaskierers besser war als bei Verwendung des Sprachmaskierers, wenn  $F_0 = 100$  Hz betrug ( $p = 0.048$ ).

**12.2 Fehlerverteilung**

Die Fehlerverteilung wird in Abbildung 12.3 und 12.4 nur für die beiden Sprachmaskierer (F0 = 100 Hz und F0 = 180 Hz) dargestellt. Die Abbildungen zeigen die Differenz der A-Typ (jeweils die ersten zwei Balken) und B-Typ-Fehler (jeweils Balken drei und vier) in Prozent für die Messungen des Kopfschattens, des Squelchs und der Summation. Aufgetragen ist der Mittelwert, berechnet über alle Probanden, für die zwei unterschiedlichen Grundfrequenzen. Die Fehlerbalken stellen das 95%ige Konfidenzintervall dar. Positive Werte zeigen eine Zunahme, negative Werte eine Verminderung der Fehler an. Ein Stern über den Säulen zeigt an, ob der jeweilige Effekt signifikant war. P-Werte können in Tabelle C.3.5 im Anhang C.3 nachgelesen werden. Signifikante Unterschiede zwischen den Fehlertypen und F0, die sich durch die 2-faktorielle ANOVA errechneten, sind unterhalb der Säulen durch einen Stern gekennzeichnet (s. Tabelle C.3.6 - C.3.8).

Bezüglich der Hörhilfen sind die Abbildungen 12.3 und 12.4 in gleicher Weise aufgebaut wie die Abbildungen 12.1 und 12.2. In Abbildung 12.3 ist der Anteil der Fehler beim Kopfschatteneffekt für das CI berechnet. Der Squelcheffekt zeigt die Veränderung des Fehleranteils für die Situation auf, in der das Störsignal von der Seite des CIs dargeboten wird und dies zum HG-Ohr hinzugeschaltet wird. Bei der Summation wird die Veränderung aufgezeigt, die die Hinzunahme des CIs zum HG bewirkt. Abbildung 12.4 ist gleichermaßen für die Effekte des HG bzw. dessen Hinzunahme aufgebaut.

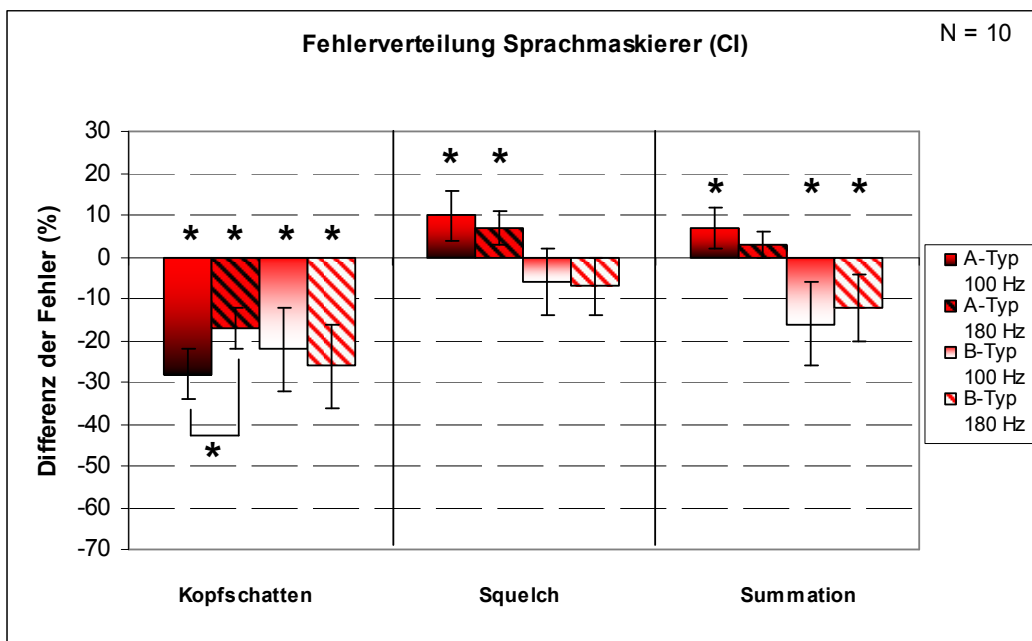


Abbildung 12.3: Differenz der A- und B-Typ-Fehler (für CI) in Prozent bei Verwendung der Sprachmaskierer mit F0 = 100 Hz und 180 Hz. Berechnungen: Kopfschatten für CI; Squelch: Störsignal auf der Seite des CIs; Summation: Hinzunahme des CIs.

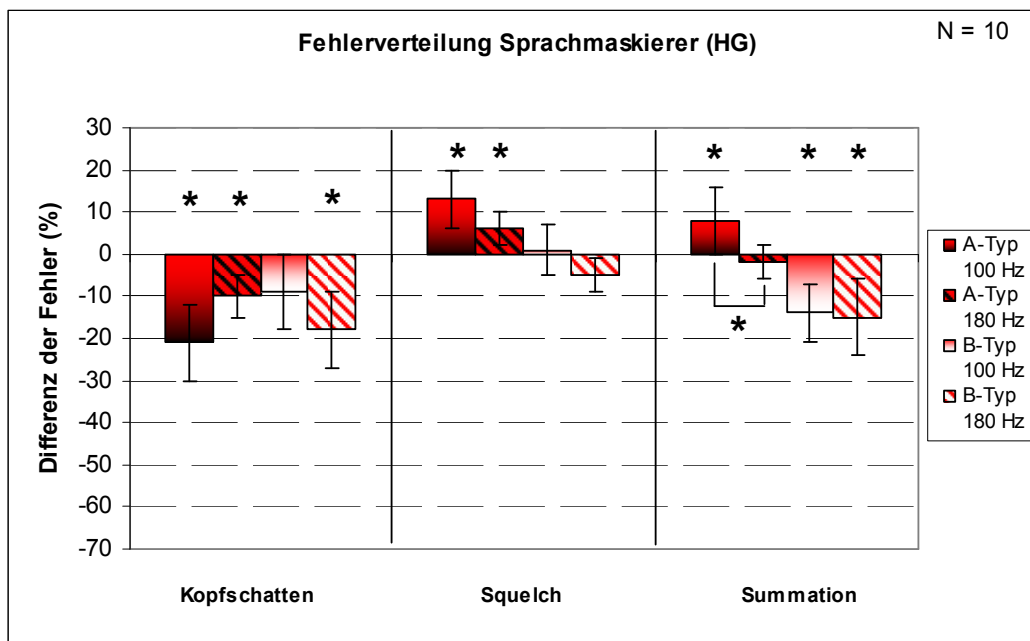


Abbildung 12.4: Differenz der A- und B-Typ-Fehler (für HG) in Prozent bei Verwendung der Sprachmaskierer mit F0 = 100 Hz und 180 Hz. Berechnungen: Kopfschatten für HG; Squelch: Störgeräusch auf Seite des HGs; Summation: Hinzunahme des HGs.

### Kopfschatten

Sowohl für das Hören mit dem CI als auch für das Hören mit dem HG wurden aufgrund des Kopfschattens die A-Typ-Fehler signifikant reduziert. Für das Hören mit dem CI wurden die B-Typ-Fehler ebenfalls signifikant reduziert. Für das Hören mit dem HG war nur der B-Typ-Fehler mit F0 = 180 Hz signifikant unterschiedlich (s. Tabelle C.3.5).

Die 2-faktorielle ANOVA zeigte einen signifikanten Unterschied für F0 in den A-Typ-Fehlern ( $p = 0.002$ ). Der anschließend durchgeführte Post-Hoc-Test zeigte, dass in der unilateralen CI-Kondition (s. Abbildung 12.3) die A-Typ-Fehler durch den Kopfschatteneffekt bei Verwendung der Sprachmaskierer bei F0 = 100 Hz stärker verringert wurden als die A-Typ-Fehler bei F0 = 180 Hz ( $p = 0.004$ ).

Für die B-Typ-Fehler ergab sich ebenfalls ein signifikanter Unterschied für F0 ( $p = 0.03$ ), der allerdings nach Bonferroni-Adjustierung keine Signifikanz mehr aufwies (s. Tabelle C.3.6).

### Squelch

Infolge des Squelchs stieg die Anzahl der A-Typ-Fehler signifikant, unabhängig davon, auf welcher Seite (Hörhilfe) das Störgeräusch präsentiert wurde. Die Anzahl der B-Typ-Fehler wurde tendenziell eher verringert. Allerdings war dieser Effekt nicht signifikant.

Nach Durchführung der ANOVA (s. Tabelle C.3.7) zeigte sich, dass F0 keinen signifikanten Einfluss auf die A-Typ-Fehler hatte. Bei Untersuchung der B-Typ-Fehler gab es einen

signifikanten Effekt für F0 (100 Hz vs. 180 Hz) ( $p = 0.046$ ). Im anschließenden Post-Hoc-Test konnte diese Signifikanz, nach Adjustierung der p-Werte, nicht nachgewiesen werden.

### Summation

Die Fehleranalyse für die binaurale Summation ergab eine signifikante Zunahme der A-Typ-Fehler in der Kondition F0 = 100 Hz, jedoch nicht in der Kondition F0 = 180 Hz. Dies galt gleichermaßen sowohl für die Hinzunahme des CIs als auch des HGs. B-Typ-Fehler wurden signifikant für beide Hörhilfen bei beiden Grundfrequenzen reduziert (s. Abbildung 12.3 und 12.4).

Mithilfe der ANOVA (s. Tabelle C.3.8) konnte ein signifikanter Unterschied für F0 bei der Analyse der A-Typ-Fehler nachgewiesen werden ( $p = 0.016$ ). Anhand der Post-Hoc-Tests zeigte sich ein Unterschied für die Kondition in der das HG dem CI zugeschaltet wurde ( $p = 0.02$ ) (s. Abbildung 12.4).

### **12.3 Diskussion**

Es wurde untersucht, inwieweit bimodal versorgte CI-Träger von Kopfschatteneffekt, binauralem Squelch und binauraler Summation in störschallerfüllten Hörsituationen profitieren. Dazu wurden zwei unterschiedliche Maskierer verwendet, Sprache und fluktuierendes Rauschen.

Die Ergebnisse zeigten, dass die Probanden nicht von allen untersuchten Größen in gleichem Maße profitierten. Dies belegten auch andere Studien (z.B. Ching et al., 2006; Schleich et al., 2004), bei denen sich durch den Kopfschatten eine verbesserte Sprachverständlichkeit ergab, für den binarualen Squelch und die binaurale Summation jedoch keine oder nur geringe Verbesserungen in der Sprachverständlichkeit beobachtet werden konnten (Gantz et al., 2002).

Die vorliegende Untersuchung geht über diese Studien hinaus, da die Effekte bei bimodal versorgten CI-Trägern erstmals für einen Sprachmaskierer untersucht wurden. Es zeigte sich, dass die untersuchten Größen Kopfschatten, Squelch und Summation bei der Verwendung von Sprachmaskierern andere Änderungen als bei der Verwendung von Rauschmaskierern erfuhren. Dies ist besonders für Messungen in der klinischen Routine interessant, da hier größtenteils Rauschmaskierer verwendet werden, um das alltägliche Sprachverstehen bei Personen mit Hörschädigung nachzuweisen. In der alltäglichen Situation kommt allerdings Sprache als Maskierer häufig vor.

Die Analyse der Fehlerverteilung bei der Verwendung der Sprachmaskierer gab einen Hinweis darauf, wodurch der Unterschied zwischen den beiden Maskierungsarten zustande kam. Nicht nur der Anteil des EM, ersichtlich anhand der falsch oder nicht verstandenen Wörter (B-Typ-Fehler), sondern ebenfalls der Anteil des IM, ersichtlich anhand der

Verwechslungen (A-Typ-Fehler), änderte sich. Durch den Kopfschatten wurden beide Fehlertypen, und damit sowohl EM als auch IM reduziert.

Bezogen auf den Einfluss von  $F_0$ , war der Kopfschatteneffekt bei Verwendung der Sprachmaskierer etwas größer für  $F_0 = 100$  Hz im Vergleich zu  $F_0 = 180$  Hz bei Nutzung des CIs. In den Rohdaten, vor Berechnung des Kopfschatteneffektes, war das Sprachverstehen in der Kondition  $F_0 = 180$  Hz besser als das Sprachverstehen in der Kondition  $F_0 = 100$  Hz. Zusätzlich wurden die Verwechslungen in der Kondition  $F_0 = 180$  Hz weniger verringert als in der Kondition  $F_0 = 100$  Hz. Nach Hawley et al. (2004) würde dies für die Effektivität der räumlichen Trennung sprechen, die ausreicht, das Sprachverstehen zu verbessern, wenn die Hörer keinen Zugriff auf weitere Merkmale haben ( $F_0 = 100$  Hz), die eine Trennung der Sprecher erlauben würden.

Hinsichtlich des Squelcheffektes verbesserte sich das Sprachverstehen bei Verwendung der Rauschmaskierer, wenn sich die Störquelle auf der Seite des CIs befand (s. Abbildung 12.1). Hingegen änderte sich das Sprachverstehen nicht, wenn sich die Störquelle auf der Seite des HGs befand (s. Abbildung 12.2). Obwohl alle Probanden mit dem HG alleine ein gutes Sprachverstehen in Ruhe aufwiesen, reichte dieses gute Sprachverstehen nicht aus, um Sprache im Störgeräusch zu verstehen. Das Sprachverstehen mit dem HG wurde sogar dann vom Sprachverstehen mit CI übertroffen, wenn das Störgeräusch ipsilateral auf der Seite des CIs angeboten wurde und somit einen schlechteren TMR aufwies (Dunn et al., 2005). Das CI dominiert auch in schlechten Hörbedingungen das Sprachverstehen (Morera et al., 2012). Das spricht dafür, dass die Probanden die Informationen des HGs zentral nicht mit denen des CIs verknüpfen konnten und weist auf ein Fehlen des Squelcheffektes und des binauralen Abgleichs hin.

Hinsichtlich des Squelcheffektes verschlechterte sich die Zielverständlichkeit bei Verwendung der Sprachmaskierer, wenn Ziel- und Maskierersprecher die gleiche Grundfrequenz (100 Hz) aufwiesen. Die Verwechslungsfehler stiegen an, wohingegen der Anteil falsch oder nicht verstandener Wörter keine deutliche Veränderung erfuhr. Bei unterschiedlicher Grundfrequenz änderte sich das Sprachverstehen aufgrund einer gegenläufigen Fehlerverteilung nicht. Der Anteil der Verwechslungen stieg an, während der Anteil falsch oder nicht verstandener Wörter verringert wurde. Dies spricht für eine Reduzierung von EM, ersichtlich anhand der Reduktion der B-Typ-Fehler, nicht jedoch für eine Reduzierung von IM (Brungart, 2001).

Der Summationseffekt trat für beide Maskierertypen auf. Dies kann einerseits durch eine Lautheitssummation erklärt werden, die bei Hinzunahme der zweiten Hörhilfe entsteht (Ching et al., 2007). Andererseits kann dies auch durch binaurale Redundanz des bereits beschriebenen Komplementäreffektes von HG und CI erklärt werden (Ching et al., 2007; Cullington and Zeng, 2011). Falsch oder nicht verstandene Wörter wurden bei Hinzunahme



## **Kapitel 12. Ergebnisse zu Experiment 3: Effekte räumlichen Hörens**

der zweiten Hörhilfe deutlich reduziert. Dieses Ergebnis unterstützt die Ergebnisse aus Experiment 2 und steht für eine Reduktion des EM (Brungart, 2001).

Bei Verwendung der Sprachmaskierer stieg der Anteil der Verwechslungen an, wenn Ziel- und Maskierersprecher die gleiche Grundfrequenz (100 Hz) hatten. Bei unterschiedlichen Grundfrequenzen (180 Hz) gab es keine Änderung. Das spricht dafür, dass die CI-Träger in der Lage waren, konkurrierende Sprecher anhand eines großen F0-Unterschiedes zu trennen, bereits vor Hinzunahme der zweiten Hörhilfe.

Besonders die Ergebnisse des Squelcheffektes lassen nicht darauf schließen, dass die Information beider Ohren zentral integriert wurde (Gantz et al., 2002). In der Hörsituation zur Berechnung des Squelcheffektes überwog das Sprachverstehen mit dem CI das Sprachverstehen mit dem HG. Garadat et al. (2009) untersuchten anhand von CI-Simulationen die Auswirkungen spektraler und binauraler Feinstruktur auf das Sprachverstehen. Um die spektrale Feinstruktur zu entfernen und die temporale Feinstruktur beizubehalten, wurden die Stimuli zuerst in Frequenzbänder (16, 8 oder 4) gefiltert und dann mit einer kopfbezogenen Übertragungsfunktion (engl. „head related transfer function, HRTF“) bearbeitet. Um sowohl die spektrale Feinstruktur als auch die temporale Feinstruktur zu entfernen, wurde dieser Prozess in umgekehrter Reihenfolge durchgeführt. In den Ergebnissen zeigte sich, dass eine geringe Anzahl von Frequenzkanälen (spektrale Feinstruktur) nicht der Grund für die Abwesenheit des Squelcheffektes sein konnte. Es zeigte sich vielmehr, dass die Abwesenheit binauraler Verarbeitung (temporale Feinstruktur) dadurch zustande kam, dass wichtige Information, wie interaurale Zeitdifferenzen (ITDs), durch zwei unabhängige Hörhilfen nicht eindeutig ausgewertet werden konnten. Diese unterschiedlichen Zeit- und Intensitätsmerkmale müssen für einen positiven Squelcheffekt ausgewertet werden können (Schafer et al., 2011).

Der Summationseffekt beruht auf redundanter Information, die durch beide Ohren zur Verfügung gestellt wird und dadurch in einem lauter wahrgenommenen Schallsignal resultiert. Durch die redundante Information werden insgesamt mehr Wörter verstanden und die falsch oder nicht verstandenen Wörter werden reduziert. Hierfür ist die Auswertung unterschiedlicher Zeit- und Intensitätsmerkmale nicht in dem Maße notwendig, wie es beim Squelcheffekt der Fall ist.

Der Unterschied im Sprachverstehen bei Verwendung eines Rauschmaskierers im Gegensatz zur Verwendung eines Sprachmaskierers kann durch Merkmale der Feinstruktur, die in Sprache, aber nicht in Rauschen vorhanden sind, erklärt werden. Die Feinstruktur kann durch akustische Verstärkung des HGs beibehalten und übertragen werden.

Insgesamt müssen Kopfschatten-, Squelch- und Summationseffekt bei CI-Trägern abhängig davon bewertet werden, ob mit Sprache oder Rauschen maskiert wird. Arbogast et al. (2002) konnten ebenfalls unterschiedliche Änderungen abhängig von der Maskiererart bei

IM und EM beobachten. Sie untersuchten in einer Simulationsstudie, ob die Fähigkeit zur räumlichen Trennung von unterschiedlichen Maskiererarten beeinflusst wird. Ohne räumliche Quellentrennung war das Zielsignal zwar wahrnehmbar, die dazugehörige Information konnte jedoch nicht vom Maskiereranteil getrennt werden. Durch die räumliche Trennung war die Separation der Sprecher möglich, so dass die Information dem jeweiligen Sprecher zugeordnet werden konnte. Die Ähnlichkeit von Ziel- und Maskierersignal wurde reduziert und führte zu einer geringeren Verwechslung beider Signale. Beim EM hingegen wurde das Verstehen eines Zielsignals eher von akustisch, physikalischen Faktoren wie dem Kopfschatteneffekt beeinflusst. Arbogast et al. (2002) erklärten, dass IM und EM auf unterschiedlichen Arten der Wahrnehmung basieren. IM beruht auf Vorgängen der zentralen Wahrnehmung und EM auf Vorgängen der peripheren Verarbeitung, so dass beide Maskiererarten von unterschiedlichen Mechanismen bei der Quellentrennung profitieren. Dies entspricht den Ergebnissen dieser Untersuchung, die im Gegensatz zu Arbogast et al. (2002) mit „realen“ CI-Nutzern gewonnen wurden.

### **12.4 Zusammenfassung der Ergebnisse**

Zusammenfassend konnte in Experiment 3 gezeigt werden, dass der Maskierertyp (Rauschen vs. Sprache) die Effekte räumlichen Hörens unterschiedlich beeinflusst. Für beide Hörhilfen (CI und HG) wurde das Sprachverstehen durch den Kopfschatteneffekt verbessert, wobei der Effekt bei Verwendung von Sprachmaskierern größer war. Dies ergab sich aus der kombinierten Reduktion von IM und EM.

Ein Squelcheffekt konnte nicht aufgezeigt werden. Die verbesserte Zielverständlichkeit bei Verwendung der Rauschmaskierer ergab sich nur durch das gute Sprachverstehen mit dem CI alleine. Bei Verwendung der Sprachmaskierer änderte sich das Sprachverstehen nicht oder verschlechterte sich sogar. Verwechslungsfehler stiegen an, falsch oder nicht verstandene Wörter wurden eher verringert.

Ein Summationseffekt ergab sich bei Verwendung beider Maskierertypen. Die Verbesserung konnte auf eine Reduzierung von EM zurückgeführt werden.

## KAPITEL 13

### Zusammenfassung und allgemeine Diskussion

Bei Personen, die mit einem Cochlea-Implantat versorgt sind, bestehen große Schwierigkeiten in Kommunikationssituationen, in denen mehrere Personen gleichzeitig sprechen. Das Ziel dieser Untersuchungen war die Beantwortung der Frage, inwieweit Cochlea-Implantat-Träger einen Sprecher in einer Kommunikationssituation mit konkurrierendem Sprecher verstehen und welche Merkmale von den Hörern genutzt werden können, um in solchen Situationen besser zu verstehen. Dies wurde vor dem Hintergrund von „Informational Masking“ und „Energetic Masking“ diskutiert. Der Fokus lag auf dem Einfluss stimmlicher Unterschiede zwischen Ziel- und Maskierersprecher, dem Einfluss unterschiedlicher Pegel und dem Einfluss unterschiedlicher räumlicher Positionen von Ziel- und Maskierersprecher. Zusätzlich zu unilateral versorgten CI-Nutzern wurden bimodal, mit einem Hörgerät auf der dem CI gegenüberliegenden Seite, versorgte Nutzer in die Studie miteinbezogen. Hier stand vor allem die zusätzliche Verarbeitung tiefer Frequenzen mit dem Hörgerät im Vordergrund. Bei den bimodal versorgten CI-Trägern stellte sich ebenfalls die Frage nach dem Zugewinn durch räumliche Trennung der konkurrierenden Signale und nach einem positiven Effekt durch binaurale Verarbeitung.

In einem Vortest wurde die Fähigkeit von CI-Trägern untersucht, inwieweit Unterschiede in der Grundfrequenz und in den Formantfrequenzen genutzt werden können, um einen Sprecher zu identifizieren und von anderen zu diskriminieren. Hypothese 1 war, dass die Sprechererkennung bei CI-Trägern funktioniert, hierbei jedoch die Grundfrequenz besser genutzt werden kann als die Formantfrequenzen. Die Ergebnisse wurden im Vergleich zu denen von Normalhörenden diskutiert. Es zeigte sich, dass CI-Träger durchaus in der Lage waren, F<sub>0</sub>-Unterschiede wahrzunehmen, dass jedoch Änderungen der Vokaltraktlänge und damit einhergehend Änderungen in den Formantfrequenzen nur schlecht oder gar nicht wahrgenommen wurden. Die Schlussfolgerung, die bzgl. alltäglicher Kommunikation gezogen wurde, war, dass CI-Nutzer bei der Identifikation eines Sprechers besonders dann Probleme haben, wenn sich die Grundfrequenz des Sprechers in einem Bereich zwischen

## **Kapitel 13. Zusammenfassung und allgemeine Diskussion**

eindeutig männlich und eindeutig weiblich befindet. Anders als die Normalhörenden können sie die Information der Formantfrequenzen, die Rückschlüsse auf Eigenschaften des Vokaltrakts und somit des Sprechers zulassen, nicht nutzen. Ebendies führt dazu, dass Sprecher mit ähnlichen Grundfrequenzen, und somit vorrangig Sprecher des gleichen Geschlechts, nur schlecht unterschieden werden können.

Die darauf folgenden Hauptexperimente zielten darauf ab, zu untersuchen, ob sich diese Ergebnisse auf eine Kommunikationssituation mit konkurrierenden Sprechern übertragen lassen. Können Grundfrequenzunterschiede ebenfalls genutzt werden, wenn Sprecher simultan präsentiert werden? Ist die Bildung auditorischer Ströme anhand der Grundfrequenz möglich?

Experiment 1 beschäftigte sich mit Änderungen der akustischen Parameter Pegel und Grundfrequenz bei unilateral versorgten CI-Trägern. Die Experimente 2 und 3 wurden mit bimodal versorgten CI-Trägern durchgeführt. Zusätzlich zu Unterschieden in der Grundfrequenz wurden Unterschiede in den Formantfrequenzen zwischen Ziel- und Maskiersignal eingefügt. Weiterhin wurde die Ausnutzung binauraler Fähigkeiten bei räumlich getrennten konkurrierenden Signalen untersucht.

Die Annahmen (H2 und H3) in Experiment 1 waren, dass Pegelunterschiede für ein besseres Sprachverstehen genutzt werden können, Grundfrequenzunterschiede hingegen nur eingeschränkt zu einem besseren Sprachverstehen beitragen. Somit sollte Informationale Masking mehr durch Pegelunterschiede als durch Grundfrequenzunterschiede reduziert werden. Die unilateral versorgten CI-Träger konnten bereits kleine Pegelunterschiede nutzen, um das Sprachverstehen zu verbessern. In Kombination mit Änderung der Grundfrequenz überwog der Einfluss der Pegelunterschiede.

Es zeigte sich, dass CI-Träger in der Lage waren, Grundfrequenzunterschiede zu nutzen, wenn mindestens ein Unterschied von 80 Hz vorlag. Bei diesem großen Unterschied wurde ein konstanter Zugewinn im Sprachverstehen bei konkurrierenden Sprechern erzielt. Große Grundfrequenzunterschiede konnten genutzt werden, um getrennte auditorische Ströme zu bilden und dem Zielsprecher zu folgen. Der Grundfrequenzunterschied führte zu einer sinkenden Anzahl von Verwechslungen. Dieses wurde in Experiment 2 bestätigt. Sowohl beim Hören mit dem CI alleine als auch in der bimodalen Kombination CI + HG stieg das Sprachverstehen bei  $\Delta F_0 = 80$  Hz an und die Verwechslungen sanken. Durch große Grundfrequenzunterschiede konnte somit Informationale Masking reduziert werden.

Dies bedeutet im Umkehrschluss, dass kleine Unterschiede in der Grundfrequenz nicht genutzt werden können, um Sprecher voneinander zu trennen. Dies entspricht den Ergebnissen aus dem Vortest, bei dem die Stimuli in Ruhe präsentiert wurden. Somit ist es

für CI-Träger schwierig, Sprecher des gleichen Geschlechts bei simultanem Auftreten, voneinander zu unterscheiden.

Unterschiede in den Formantfrequenzen konnten weder in der unilateralen CI-Kondition noch in der bimodalen CI + HG-Kondition genutzt werden, um Sprecher zu trennen. Hypothese 4, dass Grundfrequenzunterschiede besser genutzt werden können als Formantfrequenzunterschiede, wurde durch die Ergebnisse bestätigt. Dies untermauert die Aussage, dass Schwierigkeiten bei dem Verstehen mehrerer Sprecher gleichen Geschlechts entstehen.

In beiden Untersuchungen mit bimodal versorgten Patienten zeigte sich, dass das Hörgerät einen Nutzen in einer konkurrierenden Sprechersituation bringt (H5). Dieser wird dem Aspekt der Komplementarität zugesprochen. Die tieffrequente Information, übertragen durch das Hörgerät, und die höher frequente Information, übertragen durch das Cochlea-Implantat, ergänzen sich.

Die Ausnutzung vor allem der Grundfrequenz als Merkmal zur Trennung der Sprecher wurde mit dem Hörgerät nicht verbessert. Die in Hypothese 6 angenommene, zusätzliche Verminderung von Informationale Masking wurde durch ein kontralaterales Hörgerät nicht erreicht, was anhand der Analyse der Verwechslungsfehler festgemacht werden konnte. Der bimodale Zugewinn ergab sich hauptsächlich durch eine Lautheitssumation durch beide Hörhilfen und entweder eine Summation phonetischer Merkmale (Mok et al., 2006) oder eine Reduzierung von Energetic Masking. Dafür spricht die Reduzierung von falsch und nicht verstandenen Wörtern (B-Typ-Fehler). In der vorliegenden Studie ergibt sich der Zugewinn durch das Hörgerät also eher dadurch, dass insgesamt mehr Wörter (vom Ziel- und Maskierersatz) verstanden werden, anstatt dass Sprecher besser voneinander getrennt werden können. Das zusätzliche Hörgerät reduziert nur Energetic Masking, jedoch nicht Informationale Masking.

Bei den Effekten des räumlichen Hörens zeigte sich, dass Kopfschatten-, Squelch- und Summationseffekt je nach Maskierertyp unterschiedlich waren. Anhand der Fehlerverteilung konnte aufgezeigt werden, dass sich nicht nur die falsch oder nicht verstandenen Wörter, assoziiert mit Energetic Masking, sondern auch die Verwechslungsfehler, assoziiert mit Informationale Masking, änderten. Eine Reduzierung beider Maskiererarten entstand nur durch den Kopfschatteneffekt. Aufgrund des Squelch- oder Summationseffektes wurde Informationale Masking jedoch nicht reduziert. Die Ergebnisse sprechen eher für eine Reduktion des Energetic Masking.

Die drei Effekte Kopfschatten, Squelch und Summation müssen bei der Verwendung von Sprachmaskierern bei bimodal versorgten CI-Trägern anders bewertet werden als bei der Verwendung von Rauschmaskierern. Dies ist für die klinische Routine von großer

## **Kapitel 13. Zusammenfassung und allgemeine Diskussion**

Bedeutung, da bis jetzt nahezu ausschließlich Rausch- anstatt Sprachmaskierer verwendet werden, um das Sprachverstehen von CI-Trägern in alltäglichen Situationen zu bestimmen.

Für einen binauralen Zugewinn durch ein kontralaterales Hörgerät sprechen die Ergebnisse aus Experiment 3 nicht. Sowohl beim Squelcheffekt als auch bei der Summation sind es sowohl Komplementär- und Redundanzeffekte als auch die Dominanz des Sprachverstehens mit dem CI (Morera et al., 2012), die nicht auf eine zentrale Verknüpfung der Informationen, übertragen durch beide Hörhilfen, schließen lassen (H8).

Änderungen in der Grundfrequenz hatten nur wenig Einfluss auf die Effekte Kopfschatten, Squelch und Summation, wenn Nutz- und Störsignal räumlich voneinander getrennt waren.

In Experimenten mit überlagerten Signalen stellt sich die Frage nach der Fähigkeit in spektrale und / oder temporale Lücken zu hören. Normalhörende profitieren sehr von der Möglichkeit, Lücken in dem Maskiersignal zu nutzen, um auf das Nutzsignal „zugreifen“ zu können. Der „Glimpsing-Mechanismus“ ist in dieser Arbeit nicht explizit untersucht worden, es scheint aber, dass ein Lückenhören bei konkurrierenden Sprechern für die in dieser Studie getesteten CI-Träger zu einem gewissen Anteil möglich ist. Bei großen F0-Unterschieden verbesserte sich das Sprachverstehen, so dass davon ausgegangen werden kann, dass die Grundfrequenz genutzt werden konnte, um den Zielsatz als ein auditorisches Objekt zu formatieren (Shinn-Cunningham, 2008). Da die Probanden in dieser Untersuchung ein besonders gutes Sprachverstehen mit dem CI aufwiesen, sollte mit einer Verallgemeinerung für alle CI-Träger vorsichtig umgegangen werden.

Der Vorteil der räumlichen Trennung konnte nur eingeschränkt genutzt werden, um den Zielsatz zu einem auditorischen Objekt zusammenzufassen. Spricht der Nutzsprecher nicht in das „bessere“ Ohr des Patienten, steigt die Problematik, dass mehrere Sprechende verwechselt werden.

Die bimodalen CI-Nutzer in dieser Studie sind nicht vollkommen repräsentativ und vergleichbar mit Probanden anderer Studien. Die Mehrheit wies ein gutes bis sehr gutes Sprachverstehen mit dem Hörgerät auf. Das ist in der Regel nicht der Fall. Da jedoch abzusehen ist, dass das Patientenkollektiv mit ausreichend Restgehör zunehmen wird (elektro-akustische Stimulation und einseitige Schwerhörigkeit) könnten diese Ergebnisse richtungweisend für solche Patienten sein. Um auch bei der Beratung der Patienten vor der Versorgung mit einem CI eine realistische Erwartungshaltung aufzubauen, sollte nach einer neueren Studie von Morera et al. (2012) insbesondere das Hauptaugenmerk auf den Grad der Asymmetrie zwischen den beiden Ohren gelegt werden. Daraus kann sich eine Einschätzung ergeben, welche Patienten (in der bimodalen Kondition) von einem binauralen Zugewinn profitieren können und bei welchen Patienten dies eher schwierig ist.

Überträgt man die Ergebnisse dieser Studie auf den Alltag, bleiben noch einige Fragen offen. Die Bedeutung der Vorabinformation bleibt unbeantwortet. Die Probanden wussten, dass das Zielsignal immer von vorne (0°) dargeboten wird. In alltäglichen Situationen hat man in einer großen Unterhaltungsrunde ständige Sprecherwechsel. Subjektiv empfinden Patienten dies als besonders anstrengend. Auch die vorherige Eingewöhnung an die Stimme ist nicht immer gegeben. Oft wird man angesprochen oder etwas gefragt, ohne sich vorher auf den Sprechenden einstellen zu können. Das könnte das Sprachverstehen zusätzlich beeinflussen.

Zudem ist das Vokabular des OISa eingeschränkt, so dass eventuell eine höhere Mustererkennung stattfindet als in normaler Alltagssprache. Hört man z.B. aus einem Wort einen bestimmten Laut heraus oder erkennt die Länge eines Wortes, ist es möglich, einen schnelleren Abgleich im mentalen Lexikon vorzunehmen. Nur die vorkommenden Wörter werden aktiviert. Somit ist die linguistische Interferenz der Wörter des Maskierersprechers abgemildert.

Die Probanden in dieser Studie wurden allesamt mit ihren alltäglichen Einstellungen getestet. Die Mehrzahl der genutzten Kodierungsstrategien verarbeitet die Grundfrequenz nur ungenau. Nur in einigen Fällen nutzten die Probanden Strategien wie HiRes oder FSP, die die Feinstruktur eines Signals besser übertragen sollen. In dieser Arbeit wurde allerdings nicht untersucht, inwieweit ein Unterschied im Sprachverstehen der Probanden vorlag bezogen auf Aspekte wie die Kodierungsstrategie, die Stimulationsrate oder aktive Elektroden. Sicherlich haben diese technischen Gegebenheiten Einfluss auf das Sprachverstehen.

Die ungenaue Übertragung von Sprache durch die technischen Einschränkungen der CI-Verarbeitung führen dazu, dass auch kompensatorische Fähigkeiten der Hörer mit bedacht und im Rehabilitationsprozess trainiert werden sollten. Aufgrund der eingeschränkten Signalqualität entstehen fehlerhafte bottom-up Prozesse. Wiederherstellungsprozesse (top-down Prozesse), bei denen sich das Gehörte z.B. aus der Erschließung des Kontextes ergibt, sind dadurch vermindert. Bei weiteren Untersuchungen zum IM, einem zentralem Prozess, sollten sowohl physiologische als auch kognitive Aspekte miteinbezogen werden (Meister et al., 2009). Bereiche wie Aufmerksamkeit, Gedächtnis und Kompensationsmechanismen sollten insbesondere für Kommunikationssituationen im Störgeräusch auch in den Rehabilitationsprozess eingebunden werden.

## KAPITEL 14

### Ausblick

Einige offenen Fragen sollen in nachfolgenden Studien untersucht werden. Besonders die aus Experiment 3 gewonnen Erkenntnisse könnten für andersartig versorgte CI-Träger unterschiedlich ausfallen. Es bleibt aufzuzeigen, wie sich binaurale Effekte bei Verwendung eines Informationaler Maskers bei bilateral, mit zwei CIs, versorgten Patienten auswirken und ob auch hier ein Unterschied bei Verwendung eines Rausch- bzw. eines Sprachmaskierers existiert. Da die Feinstruktur mit zwei CIs nicht übertragen wird, kann angenommen werden, dass Rausch- und Sprachmaskierer ähnliche Ergebnisse liefern.

Die geänderten Indikationskriterien führen auch zu dem Probandenkollektiv der immer größer werdenden Gruppe der einseitig normalhörenden und kontralateral mit einem CI versorgten Patienten. Besonders in Hinblick auf die Verarbeitung unterschiedlicher Maskierertypen sollte geschaut werden, welchen Einfluss das CI auf das normalhörende Ohr in Situationen hat, in denen die binaurale Verarbeitung eine Verbesserung bringen könnte. Werden modulierte Störsignale und stationäre Störsignale mit beiden Ohren (NH + CI) eher so verarbeitet, wie man es von einem normalhörenden Ohr kennt oder wie man es von der Verarbeitung mit einem CI kennt? Das Hören in modulierten Störgeräuschen ist für NH aufgrund des Lückenhörens einfacher, für CI-Träger hingegen aufgrund der eingeschränkten Verarbeitung des CIs schwieriger.

Bei Verwendung von Sprache als Störsignal erschweren sowohl die Überlagerung spektraler Energie (Energetic Masking) als auch die zusätzliche irrelevante Information (Informational Masking) das Verstehen eines Nutzsignals. D.h. beide Arten von Maskierung sind bei Sprache als Störgeräusch vorhanden. Bisher ist unklar, wie hoch der einzelne Beitrag von EM und IM bei Verwendung eines Sprachmaskierers ist. Es stellt sich die Frage, ob sich die Effekte eventuell aufsummieren oder eine andere Verteilung vorliegt. Anhand der Ergebnisse von Experiment 3 zeigte sich, dass die Verwendung von sprachähnlichen Rauschmaskierern nicht ausreicht, um das Sprachverstehen bei der Verwendung von



Sprachmaskierern vorherzusagen. Die Ergebnisse fallen für beide Arten von Maskierern unterschiedlich aus. In der klinischen Routine werden Sprachverständlichkeitsmessungen bei CI-Trägern mit modulierten Rauschsignalen durchgeführt, in der Annahme, sprachliche Eigenschaften zu simulieren.

In einer Studie von Agus et al. (2009) wurde mit Normalhörenden (NH) ein „Speech-in-speech-in-noise“-Paradigma getestet, durch das die Effekte von EM und IM getrennt voneinander betrachtet werden können. Dieses Paradigma ermöglichte es, die auf EM basierende Verschlechterung der Zielsatzverständlichkeit vorherzusagen und anhand von Sprachverständlichkeitsmessungen zu verifizieren. Die über den vorhergesagten Anteil hinausgehende Einschränkung der Sprachverständlichkeit wurde IM zugeschrieben werden. Ziel ist es, dieses Paradigma bei CI-Trägern anzuwenden, um zu untersuchen, in welcher Weise die Effekte von EM und IM das Sprachverstehen von CI-Trägern im Vergleich zu Normalhörenden (NH) beeinflussen. Es wird die Hypothese aufgestellt, dass die auf EM basierende Verschlechterung der Zielverständlichkeit nicht dem vorhergesagten Anteil entspricht, da die Modulation von Sprache einen weiteren Störfaktor darstellt. Zusätzliches IM sollte einen geringeren Einfluss auf das Sprachverstehen haben als bei NH (Pyschny et al., 2009), da die Modulation bereits wie eine Art IM wirkt. Es bleibt aufzuzeigen, wieviel Einfluss die Information des Gesagten hat.

Weiterhin kann hier auch das Problem des „Lückenhörens“ näher untersucht werden. Bei Generierung der Stimuli ist die Grundlage immer ein stationäres Rauschen, das entweder mit einem modulierten Rauschsignal oder mit einem Sprachsignal in unterschiedlichen Signal-Rausch Abständen zusammengemischt wird. Mit jeweils einem dieser Störsignale wird der Zielsprecher überlagert. Das stationäre Rauschen füllt somit die „Lücken“, die bei einem modulierten Signal genutzt werden können, um den Zielsprecher zu verstehen.

Die bis dato im deutschsprachigen Raum verwendeten Wort- und Satztests werden ausschließlich von Männerstimmen gesprochen. Es fehlt somit die Repräsentation von Frauen- und Kinderstimmen. Auch im Rahmen von Untersuchung zur Sprachverständlichkeit im Störgeräusch bzw. mit konkurrierenden Sprechern wird in bisherigen Studien in den meisten Fällen eine Männerstimme als Zielsprecher verwendet. Besonders in Studien mit konkurrierenden Sprechern liegt das Hauptaugenmerk eher auf der Diskriminationsfähigkeit der Probanden, also inwieweit überlagerte Sprecher getrennt werden können und welcher stimmliche akustische Unterschied dies zulässt. Nur wenige Studien haben sich mit der grundlegenden Verständlichkeit des Zielsprechers beschäftigt. Die Frage, inwieweit sich das Sprachverstehen verändern könnte, wenn gleiches Sprachmaterial von unterschiedlichen Stimmen gesprochen würde, ist ebenfalls noch ungeklärt. Das Ziel einer weiterführenden

Studie wäre zu untersuchen, ob unterschiedliche Zielsprecher einen Einfluss auf das Sprachverstehen von CI-Trägern besonders im Störgeräusch haben.

Ergebnisse von Studien (Darwin et al., 2003; Mackersie et al., 2011) mit NH und hörgeschädigten Probanden deuteten darauf hin, dass diese Probandengruppen von einem wachsenden Unterschied der beiden überlagerten Sprecherstimmen profitieren, wenn der Zielsprecher eine höhere Grundfrequenz hat. Eine Verbesserung durch einen zunehmenden Unterschied zwischen den Sprecherstimmen zeigte sich nicht, wenn der Zielsprecher eine tiefe Grundfrequenz hatte.

Ergebnisse aus Studien mit CI-Trägern sind eher gegensätzlich zu denen der NH. Loizou et al. (2008) fanden heraus, dass die Vokalidentifikation bei CI-Trägern mit zunehmender F0 der Sprecher sinkt. Männerstimmen konnten verglichen mit Frauen- und Kinderstimmen am besten verstanden werden. Eine Erklärung dafür sind die in CIs immer breiter werdenden Kanäle im hochfrequenten Bereich. Je höher F0, desto höher sind auch die Formanten F1 und F2, die zur Vokalidentifikation benötigt werden. Bei höheren Stimmen sind die Amplituden in den hochfrequenten Kanälen stärker. Da die Kanäle in diesem Bereich allerdings relativ breit sind, wird die Identifikation schlechter.

Hong & Turner (2009) untersuchten die Fähigkeit von CI-Trägern, Periodizitätsmerkmale zu nutzen, um unterschiedliche Sequenzen amplitudenmodulierter Geräusche zu „streamen“. Die Ergebnisse deuteten darauf hin, dass CI-Träger nur bei niedrigen Modulationsfrequenzen (ca. 80 Hz amplitudenmoduliert) Sequenzen „streamen“ konnten. Im Hinblick auf das Sprachverstehen könnte dieses Ergebnis implizieren, dass CI-Nutzer bei Sprechern mit niedriger F0 Periodizitätsmerkmale nutzen können, die dabei helfen, zwei Sprecher voneinander zu trennen und den Zielsprecher zu verstehen.

## Literaturverzeichnis

- Anderson, I., Baumgartner, W.D., Böheim, K., Nahler, A., Arnolder, C. und D'Haese, P. (2006). Telephone use: What benefit do cochlear implant users receive? *International Journal of Audiology*, 45 (8), 446 - 453.
- Arbeitsgemeinschaft deutschsprachiger Audiologen und Neurootologen und Otolologen (ADANO) (2012). AWMF-Leitlinien zur Cochlea-Implantat Versorgung einschließlich zentral-auditorischer Implantate.  
URL <http://www.hno.org/adano/leitlinien.htm> (letzter Zugriff: 26.11.2012)
- Arbogast, T.L., Mason, C.R. und Kidd Jr., G. (2002). The effect of spatial separation on informational and energetic masking of speech. *Journal of the Acoustical Society of America*, 112 (5), 2086 - 2097.
- Arbogast, T.L., Mason, C.R. und Kidd Jr., G. (2005). The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 117 (4), 2169 - 2180.
- Arndt, S., Aschendorff, A., Laszig, R., Beck, R., Schild, C., Kroeger, S., Ihorst, G. und Wesarg, T. (2010). Comparison of pseudobinaural hearing to real binaural hearing rehabilitation after cochlear implantation in patients with unilateral deafness and tinnitus. *Otology and Neurotology*, 32 (1), 39 - 47.
- Assmann, P. F. und Summerfield, Q. A. (1990). Modelling the perception of concurrent vowels: Vowels with different fundamental frequencies. *Journal of the Acoustical Society of America*, 88 (2), 680 - 697.
- Başkent, D. (2012). Effect of speech degradation on top-down repair: Phonemic restoration with simulations of cochlear implants and combined electric-acoustic stimulation. *Journal of the Association for research in Otolaryngology*, 13 (5), 683 - 692.
- Baumann, U. und Helbig, S. (2009). Hören mit kombinierter elektrischer und akustischer Stimulation. *HNO*, 57 (6), 542 - 550.
- Blamey, P., Artieres, F., Başkent, D., et al. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: an update with 2251 patients. *Audiology and Neurotology*, 18 (1), 36 - 47.

- Binns, C. und Culling, J. F. (2007). The role of fundamental frequency contours in the perception of speech against interfering speech. *Journal of the Acoustical Society of America*, 122 (3), 1765 - 1776.
- Boersma, P. und Weenink, D. (1996). Praat, a system for doing Phonetics by Computer, version 4.4.22. Institute of Phonetic Sciences, University of Amsterdam; [www.praat.org](http://www.praat.org).
- Böhme, G. und Welzl-Müller, K. (2005). *Audiometrie – Hörprüfungen im Erwachsenen- und Kindesalter* (5. Auflage). Bern: Huber Verlag.
- Bolia, R.S., Nelson, W.T., Ericson, M.A. und Simpson, B.D. (2000). A speech corpus for multitalker communications research. *Journal of the Acoustical Society of America*, 107 (2), 1065 - 1066.
- Brand, T. und Hohmann, V. (2002). An adaptive procedure for categorical loudness scaling. *Journal of the Acoustical Society of America*, 112 (4), 1597 - 1604.
- Bregman, A.S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.
- Bronkhorst, A.W. und Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *Journal of the Acoustical Society of America*, 83 (4), 1506 - 1508.
- Bronkhorst, A.W. und Plomp, R. (1990) A clinical test for the assessment of binaural speech perception in noise. *Audiology*, 29 (5), 275 - 285.
- Brokx, J.P.L. und Nootboom, S.G. (1982). Intonation and the perceptual separation of simultaneous voices. *Journal of Phonetics*, 10, 23 - 36.
- Brown, C.A. und Bacon, S.P. (2009). Low-frequency speech cues and simulated electric acoustic hearing. *Journal of the Acoustical Society of America*, 125 (3), 1658–1665.
- Brungart, D.S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *Journal of the Acoustical Society of America*, 109 (3), 1101 - 1109.
- Brungart, D.S., Simpson, B.D., Ericson, M.A. und Scott, K.R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *Journal of the Acoustical Society of America*, 110 (5), 2527 - 2538.
- Brungart, D.S. und Simpson, B.D. (2002). Within-ear and across-ear interference in the cocktail-party listening task. *Journal of the Acoustical Society of America*, 112 (6), 2985 - 2995.
- Brungart, D.S., Ericson, M.A. und Simpson, B.D. (2002). Design considerations for improving effectiveness of multitalker speech displays. *Proceedings of the 2002 International conference on auditory displa*, Kyoto, 1 - 7.
- Büchler, M., Waikong, L. und Dillier, N. (2009). Musikwahrnehmung bei bimodaler Stimulation. *Zeitschrift für Audiologie*, 48 (1), 16 - 24.

- Buss, E., Pillsbury, H., Buchman, C., Pillsbury, C. und Clark, M. (2008). Multicenter U.S. bilateral Med-El cochlear implantation study: Speech perception over the first year of use. *Ear and Hearing*, 29 (1), 20 - 32.
- Carhart, R. (1965). Problems in the measurement of speech discrimination. *Archives of Otolaryngology*, 82, 253 - 260.
- Carhart, R., Tillman, T.W. und Johnson, K.R. (1968). Effects of interaural time delays on masking by two competing signals. *Journal of the Acoustical Society of America*, 43 (6), 1223 - 1230.
- Carhart, R., Tillman, T.W. und Greetis, E.S. (1969). Perceptual masking in multiple sound backgrounds. *Journal of the Acoustical Society of America*, 45 (3), 694 - 703.
- Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 25 (5), 975 - 979.
- Ching T.Y.C., Incerti, P., und Hill, M. (2004). Binaural benefits for adults who use a hearing aid and a cochlear implant in opposite ears. *Ear and Hearing*, 25 (1), 9 - 21.
- Ching, T.Y.C., Incerti, P., Hill, M. und v. Wanrooy, E. (2006). An overview of binaural advantages for children and adults who use binaural/bimodal hearing devices. *Audiology and Neurotology*, 11, 6 - 11.
- Ching, T.Y.C., v. Wanrooy, E. und Dillon, H. (2007). Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: A review. *Trends in Amplification*, 11 (3), 161 - 192.
- Cochlear Deutschland GmbH & Co. KG; [www.cochlear.com/de](http://www.cochlear.com/de) (letzter Zugriff:02.04.2012).
- Cooke, M. (2005). Making sense of everyday speech: A glimpsing account. In: Divenyi P. (Ed.). *Speech separation by humans and machines*. Norwell, Klüwer Academic Publishers, 305 - 314.
- Cool Edit Pro (Version 2.0). Syntrillium Software Corporation, P.O. Box 62255, Phoenix, 85082.
- Crystal, D. (1998). *Die Cambridge Enzyklopädie der Sprache*. Köln: Parkland Verlag.
- Cullington, H.E. und Zeng, F.G. (2008). Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects. *Journal of the Acoustical Society of America*, 123 (1), 450 - 461.
- Cullington, H.E. und Zeng, F.G. (2011). Comparison of bimodal and bilateral cochlear implant users on speech recognition with competing talker, music perception, affective prosody discrimination, and talker identification. *Ear and Hearing*, 32 (1), 16 - 30.
- Darwin, C.J., Brungart, D.S. und Simpson, B.D. (2003). Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. *Journal of the Acoustical Society of America*, 114 (5), 2913 - 2922.

- Drennan, W.R. und Rubinstein, T.R. (2008). Music perception in cochlear implant users and its relationship with psychophysical capabilities. *Journal of the Rehabilitation of Research and Development*, 45 (5), 779 - 789.
- Dunn, C.C., Tyler, R.S. und Witt, S.A. (2005). Benefit of wearing a hearing aid on the unimplanted ear in adult users of a cochlear implant. *Journal of Speech Language and Hearing Research*, 48 (3), 668 - 680.
- Durlach, N.I. und Colburg, H.S. (1978). Binaural phenomena. In E.C. Carterette & M.P. Friedman (Eds.). *Handbook of perception, Volume IV*, New York: Academic Press: 365 - 466.
- Durlach, N.I., Mason, C.R., Shinn-Cunningham, B.C., Arbogast, T.L., Colburn, H.S. und Kidd Jr., G. (2003). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *Journal of the Acoustical Society of America*, 114 (1), 368 - 379.
- Eckert, H. und Laver, J. (1994). *Menschen und ihre Stimmen - Aspekte der vokalen Kommunikation*. Weinheim: Beltz Verlag.
- Festen, J.M. und Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, 88 (4), 1725 - 1736.
- Fletcher, H. (1929). *Speech and hearing*. New York: Van Nostrand.
- Fletcher, H. (1940). Auditory patterns. *Reviews of modern physics*, 12, 47 - 65.
- Freyman, R.L., Helfer, K.S., McCall, D.D. und Clifton, R.K. (1999). The role of perceived spatial separation in the unmasking of speech. *Journal of the Acoustical Society of America*, 106 (6), 3578 - 3588.
- Freyman, R.L., Balakrishnan, U. und Helfer, K.S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *Journal of the Acoustical Society of America*, 115 (5), 2246 - 2256.
- Fu, Q.-J., Chinchilla, S., Nogaki, G. und Galvin, J.J. (2005). Voice gender identification by cochlear implant users: The role of spectral and temporal resolution. *Journal of the Acoustical Society of America*, 118 (3), 1711 - 1718.
- Gantz, B., Tyler, R.S., Rubinstein, J.T., Wolaver, A., Lowder, M., Abbas, P., Brown, C., Hughes, M. und Preece, J.P. (2002). Binaural cochlear implants placed during the same operation. *Otology and Neurotology*, 23 (2), 169 - 180.
- Garadat, S.N., Litovsky, R. Yu, G. und Zeng, F.G. (2009). Role of binaural hearing in speech intelligibility and spatial release from masking using vocoded speech. *Journal of the Acoustical Society of America*, 126 (5), 2522 - 2535.

- Gaudrain, E., Grimault, N., Healy, E.W. und Béra, J.-C. (2008). Streaming of vowel sequences based on fundamental frequency in a cochlear-implant simulation. *Journal of the Acoustical Society of America*, 124 (5), 3076 - 3087.
- Diller, G. (2009): Hörschäden. In Grohnfeldt, M. (Hrsg). *Lehrbuch der Sprachheilpädagogik und Logopädie. Diagnostik, Prävention und Evaluation. Band 3. 3. Auflage.* Stuttgart: W. Kohlhammer GmbH
- Hansen, M (2007). *Psychoakustik I, Vorlesung Hörakustik und Audiologie -- 1161 Version 1.4*, Oldenburg.
- Hawley, M., Litovsky, R. und Culling, J. (2004). The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *Journal of the Acoustical Society of America*, 115 (2), 833 - 843.
- Hellbrück, J. (2004). *Hören: Physiologie, Psychologie und Pathologie*. Göttingen: Hofgrefe.
- Helms, J. und Müller, J. (1999). Die Auswahl eines Cochlea-Implants und die Ergebnisse der Implantation. *Laryngology-Rhinology-Otolaryngology*, 78, 12 - 13.
- Hess, W. (2003). *Grundlagen der Phonetik – 4. Suprasegmentale Merkmale und Prosodie*. Vorlesung Universität Bonn. Version vom 23.01.2003
- Hessel, H. (2004). *High-rate stimulation in Cochlea-Implantaten*. Köln: Vortrag
- Hessel, H. und Hey, M. (2004). *Aspekte der ein- und zweiseitigen CI-Versorgung*. Hannover, Cochlear Deutschland GmbH & Co. KG.
- Hong, R.S. und Turner, C.W. (2009). Sequential stream segregation using temporal periodicity cues in cochlear implant recipients. *Journal of the Acoustical Society of America*, 126 (1), 291 - 299.
- Hoth, S. (1999). Medizinische Akustik und Audiologie. In Bille, J. und W. Schlegel (Hrsg.) *Medizinische Physik 1*. Berlin Heidelberg: Springer Verlag: 305 - 426.
- Hoth, S. und Müller-Deile, J. (2009). Audiologische Rehabilitation von Cochleaimplantat Trägern. *HNO*, 57 (7), 635 - 648.
- Jacob, R., Stelzig, Y., Nopp, P. und Schleich, P. (2011). Audiologische Ergebnisse mit Cochleaimplantat bei einseitiger Taubheit. *HNO*, 59 (5), 453 - 460.
- Jones, J.A. und Freyman, R.L. (2012). Effect of priming on energetic and informational masking in a same-different task. *Ear and Hearing*, 33 (1), 124 - 133.
- Kießling, J., Kollmeier, B. und Diller, G. (2008). *Versorgung und Rehabilitation mit Hörgeräten (2. Auflage)*. Stuttgart: Thieme Verlag KG.
- Kidd, Jr. G., Mason, C.R., Deliwal, P.S., Woods, W.S. und Colburn, H.S. (1994). Reducing informational masking by sound segregation. *Journal of the Acoustical Society of America*, 95 (6), 3475 - 3480.

- Kidd, Jr. G., Mason, C.R., Arbogast, T.L., Brungart, D.S. und Simpson, B.D. (2003). Informational masking caused by contralateral stimulation. *Journal of the Acoustical Society of America*, 113 (3), 1594 - 1603.
- Kidd, Jr. G., Arbogast, T.L., Mason, C.R. und Gallun, F.J. (2005). The advantage of knowing where to listen. *Journal of the Acoustical Society of America*, 118 (6), 3804 - 3815.
- Kidd, Jr. G., Mason, C.R., Richards, V.M., Gallun, F.J. und Durlach, N.I. (2008). Informational Masking. In Yost, W. A., A.N. Popper und R.R Fay (Hrsg.): *Auditory Perception of Sound Sources*, New York: Springer, 143 - 189.
- Kong, Y.-Y., Cruz, R., Jones, J.A. und Zeng, F.G. (2004). Music perception with temporal cues in acoustic and electric hearing. *Ear and Hearing*, 25 (2), 173 - 185.
- Kong, Y.-Y., Stickney, G.S. und Zeng, F.-G. (2005). Speech and melody recognition in binaurally combined acoustic and electric hearing. *Journal of the Acoustical Society of America*, 117 (3), 1351 - 1361.
- Kong, Y.-Y. und Carlyon, R.P. (2007). Improved speech recognition in noise in simulated binaurally combined acoustic and electric stimulation. *Journal of the Acoustical Society of America*, 121 (6), 3717 - 3727.
- Kong, Y.-Y., Mullangi, A., Marozeau J. und Epstein, M. (2011). Temporal and spectral cues for musical timbre perception in electric hearing. *Journal of Speech, Language and Hearing Research*, 54 (3), 981 - 994.
- Kong, Y.-Y., Mullangi, A. und Marozeau, J. (2012). Timbre and speech perception in bimodal and bilateral cochlear-implant listeners. *Ear and Hearing*, 33 (5), 645 - 659.
- Kovacic, D. und Balaban, E. (2009). Voice gender perception by cochlear implantees. *Journal of the Acoustical Society of America*, 126 (2), 762 - 775.
- Laszig, R., Aschendorff, A., Schipper, J. und Klenzner, T. (2004a). Aktuelle Entwicklung zum Kochleaimplant. *HNO-Praxis*, 4, 357 - 362.
- Laszig, R., Aschendorff, A., Stecker, M., Müller-Deile, J., Maune, S., Dillier, N., Weber, B., Hey, M., Begall, K., Lenarz, T., Battmer, R.D., Böhm, M. et al. (2004b). Benefits of bilateral electrical stimulation with the nucleus cochlear implant in adults: 6 months postoperative results. *Otolaryngology and Neurotology*, 25, 958 - 968.
- Leek, M.R., Brown, M.E. und Dorman, M.F. (1991). Informational Masking and auditory attention. *Perception and Psychophysics*, 50 (3), 205 - 214.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.
- Lenarz, T., Hartrampf, R., Battmer, R.D., Bertram, B. und Lesinski, A. (1996). Cochlear Implantation bei jungen Kindern. *Laryngorhinootology*, 75, 719 - 726.
- Lenarz, T. und Boenninghaus, H.-G. (2012). *HNO (14. Auflage)*. Berlin Heidelberg: Springer Verlag.



- Li, L., Daneman, M., Qi, J.G. und Schneider, B.A. (2004). Does the information content of an irrelevant source differentially affect spoken word recognition in younger and older adults? *Journal of Experimental Psychology: Human Perception and Performance*, 30 (6), 1077 - 1091.
- Li, N. und Loizou, P.C. (2007). Factors influencing glimpsing of speech in noise. *Journal of the Acoustical Society of America*, 122 (2), 1165 - 1172.
- Li, N. und Loizou, P.C. (2008). A glimpsing account for the benefit of simulated combined acoustic and electric hearing. *Journal of the Acoustical Society of America*, 123 (4), 2287 - 2294.
- Loizou, P.C. (1998). Mimicking the human ear: An overview of signal-processing strategies for converting sound into electrical signals in cochlear implants. *IEEE Signal Processing Magazine September*, 101 - 130.
- Loizou, P.C., Hu, Y., Litovsky, R., Yu, G., Peters, R., Lake, J. und Roland, P. (2009) Speech recognition by bilateral cochlear implant users in a cocktail-party setting. *Journal of the Acoustical Society of America*, 125 (1), 372 - 383.
- Mackersie, C. L., Dewey, J. und Guthrie, L.A. (2011). Effect of fundamental frequency and vocal-tract length cues on sentence segregation by listeners with hearing loss. *Journal of the Acoustical Society of America*, 130 (2), 1006 - 1019.
- Maurer, J. (2009). Gegenwärtiger Stand der Kochleaimplantatversorgung bei Erwachsenen und Kindern. *HNO*, 57 (7), 693 - 702.
- McDermott, H. J. (2004). Music Perception with Cochlear Implants: A Review. *Trends in Amplification*, 8 (2), 49 - 82.
- Med-El Elektromedizinische Geräte GmbH; [www.medel.com](http://www.medel.com) (letzter Zugriff:02.04.2012).
- Med-El Elektromedizinische Geräte GmbH (2008). *Cochlea – Implantat verstehen*. Innsbruck.
- Med-El Elektromedizinische Geräte GmbH (2011). Focus on: FineHearing™ Technologie - A step closer to natural hearing.  
[URL: http://www.medel.com/data/downloads/MAESTRO\\_US/FocusOn\\_FineHearing.pdf](http://www.medel.com/data/downloads/MAESTRO_US/FocusOn_FineHearing.pdf)  
(letzter Zugriff 22.03.2013)
- Meister, H., Pyschny, V., Landwehr, M., Wagner, P., Walger, M., v. Wedel, H. (2008) Konzeption und Realisierung einer Prosodie-Testbatterie. *HNO*, 56 (3), 340 - 348.
- Meister, H., Landwehr, M., Pyschny, V., Walger, M. und v. Wedel, H. (2009). The perception of prosody and speaker gender in normal-hearing listeners and cochlear implant recipients. *International Journal of Audiology*, 48 (1), 38 - 48.
- Meister, H., Landwehr, M., Pyschny, V., Wagner, P. und Walger, M. (2011a). The perception of sentence stress in cochlear implant recipients. *Ear and Hearing*, 32 (4), 459 - 467.

- Meister, H., Landwehr, M., Pyschny, V., Grugel, L. und Walger, M. (2011b). Use of intonation contours for speech recognition in noise by cochlear implant recipients. *Journal of the Acoustical Society of America*, 129 (5), 204 - 209.
- Mitterbacher, A. (2004). *Neue Signalverarbeitungs- und Stimulationsstrategien für Cochlea-Implantate*. Dissertation, Technische Universität Wien.
- Mok, M., Grayden, D., Dowell, R.C. und Lawrence, D. (2006). Speech perception for adults who use hearing aids in conjunction with cochlear implants in opposite ears. *Journal of Speech Language and Hearing Research*, 49 (2), 338 - 351.
- Moore, T. (1981). Voice communication jamming research. In *AGARD Conference Proceedings 331. Aural Communication in Aviation*, Neuilly-Sur-Seine, France, 2:1 - 2:6.
- Moore, B.C.J. (2003). *An introduction to the psychology of hearing (5<sup>th</sup> edition)*. New York: Academic Press.
- Morera, C., Manrique, M., Ramos, A., et al. (2005). Advantages of binaural hearing provided through bimodal stimulation via a cochlear implant and a conventional hearing aid: A 6-month comparative study. *Acta Oto-Laryngologica*, 125 (6), 596 - 606.
- Morera, C., Cavalle, L., Manrique, M., et al. (2012). Contralateral hearing aid use in cochlear implanted patients: multicenter study of bimodal benefit. *Acta Oto-Laryngologica*, 132 (10), 1084 - 1094.
- Müller, J., Schön, F. und Helms, J. (2002). Speech understanding in quiet and noise in bilateral users of the MED-EL COMBI 40/40+ cochlear implant system. *Ear and Hearing*, 23 (3), 198 - 206.
- Müller-Deile, J (2008). Evaluation binauralen Hörens mit Cochlea-Implantaten. 39. *DGMP Tagung*: Oldenburg.
- Müller-Deile, J. (2009). *Verfahren zur Anpassung und Evaluation von Cochlear Implant Sprachprozessoren*. Heidelberg: Median Verlag.
- Neff, D. L. und Green, D.M. (1987). Masking produced by spectral uncertainty with multicomponent maskers. *Perception and Psychophysics*, 41 (5), 409 - 415.
- Nelson, P.B., Jin, S.H., Carney, A.E. und Nelson, D.A. (2003). Understanding speech in modulated interference: cochlear implant users and normal-hearing listeners. *Journal of the Acoustical Society of America*, 113 (2), 961 - 968.
- Nelson, P.B. und Jin, S.H. (2004). Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners. *Journal of the Acoustical Society of America*, 115, 2286 - 2294.
- Netaudio (2003). *Auditive Wahrnehmung und kritische Bandbreiten*. Berlin  
URL: <http://www.netaudio.de/psychoakustik/> (letzter Zugriff 09.02.2013)
- Nie, K., Barco, A. und Zeng, F.-G. (2006). Spectral and temporal cues in cochlear implant speech perception. *Ear and Hearing*, 27 (2), 208 - 217.

- Optik Lau Akustik. [www.optik-lau-akustik.de](http://www.optik-lau-akustik.de) (letzter Zugriff 09.02.2013)
- Peterson, G.H. und Barney, H.L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 24 (2), 175 - 184.
- Pollack, I. (1975). Auditory informational masking. *Journal of the Acoustical Society of America*, 57 (1), 5.
- Pompino-Marschall, B. (1995). *Einführung in die Phonetik*. Berlin: De Gruyter & Co.
- Probst, R., Grevers, G. und Iro, H. (2008). *Hals-Nasen-Ohren-Heilkunde (3. Auflage)*. Stuttgart: Thieme Verlag KG.
- Pyschny, V., Landwehr, M., Walger, M., v. Wedel, H. und Meister, H. (2009). Untersuchungen zum „Informational Masking“ bei Kochleaimplantat-Nutzern. HNO, 57 (7), 671 - 677.**
- Pyschny, V., Landwehr, M., Hahn, M., Walger, M., v. Wedel, H. und Meister, H. (2011). Bimodal hearing and speech perception with a competing talker. Journal of Speech Language and Hearing Research, 54 (5), 1400 - 1415.**
- Qin, M.K. und Oxenham, A.J. (2003). Effect of simulated cochlear-implant processing on speech reception in fluctuating maskers. *Journal of the Acoustical Society of America*, 114 (1), 446 - 454.
- Rubinstein, J.T. (2004). How cochlear implants encode speech. *Current Opinion in Otolaryngology and Head and Neck Surgery*, 12, 444 - 448.
- Schafer, E.C., Amlani, A.M., Paiva, D., Nozari, L. und Verret, S. (2011). A meta-analysis to compare speech recognition in noise with bilateral cochlear implants and bimodal stimulation. *International Journal of Audiology*, 50 (12), 871 - 880.
- Schleich, P., Nopp, P. und D'Haese, P. (2004). Head shadow, squelch, and summation effects in bilateral users of the Med-El Combi 40/40+ cochlear implant. *Ear and Hearing*, 25 (3), 197 - 204.
- Shinn-Cunningham, B.G. und Best, V. (2008). Selective attention in normal and impaired hearing. *Trends in Amplification*, 12 (4), 283 - 299.
- Skinner, M.W. (2003). Optimizing cochlear implant speech performance. *The Annals of Otolaryngology, Rhinology and Laryngology. Supplement*, 191, 4 - 13.
- Smith, Z.M., Delgutte, B. und Oxenham, A.J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416, 87 - 90.
- SPSS Inc., IBM Company Headquarters, 233 S. Wacker Drive, 11th floor, Chicago, Illinois 60606.
- Steffens, T. (2009). Wo helfen Hörgeräte tatsächlich und wo sind ihre Grenzen? *Schnecke*, 65, 18 - 21.
- Stickney, G., Litovsky, R. und Assmann, P. (2004). Cochlear implant speech recognition with speech maskers. *Journal of the Acoustical Society of America*, 116 (2), 1081 - 1091.

- Stickney, G.S., Loizou, P.C., Mishra, L.N., Assmann, P.F., Shannon, R.V. und Opie, J.M. (2006). Effects of electrode design and configuration on channel interactions. *Hearing research*, 211, 33 - 45.
- Tillman, T.W., Kasten, R.N. und Horner, I.S. (1963): Effect of head shadow on reception of speech. *American Speech-Language Hearing Association Journals*, 5, 778 - 779.
- Trinh, D. Mazhayskaya, A. und Paterko, L. (2006). *Vokale*.  
URL: [https:// www.cl.uni-heidelberg.de/courses/archiv/ss06/ap/vortrag/Vokale.pdf](https://www.cl.uni-heidelberg.de/courses/archiv/ss06/ap/vortrag/Vokale.pdf)  
(letzter Zugriff 04.2007).
- Turner, C.W., Gantz, B.J., Vidal, C., Behrens, A. und Henry B.A. (2004). Speech recognition in noise for cochlear implant listeners: Benefits of residual acoustic hearing. *Journal of the Acoustical Society of America*, 115 (4), 1729 - 1735.
- Tyler, R.S., Parkinson, A.J., Wilson, B.S., Witt, S., Preece, J.P. und Noble, P. (2002). Patients utilizing a hearing aid and a cochlear implant: Speech perception and localization. *Ear and Hearing*, 23 (2), 98 - 105.
- Ulrich, J. (2006). Binaurales und räumliches Hören. *Hörakustik*, 9, 108 - 110.
- van de Heyning, P., Vermeire, K., Diebl, M., Nopp, P., Anderson, I. und de Ridder, D. (2008). Incapacitating unilateral tinnitus in single-sided deafness treated by cochlear implantation. *The Annals of Otology, Rhinology and Laryngology*, 117, 645 - 652.
- van Hoesel, R. und Tyler, R. (2003). Speech perception, localization, and lateralization with bilateral cochlear implants. *Journal of the Acoustical Society of America*, 113 (3), 1617 - 1630.
- Vongphoe, M. und Zeng, F.-G. (2005). Speaker recognition with temporal cues in acoustic and electric hearing. *Journal of the Acoustical Society of America*, 118 (2), 1055 - 1061.
- von Ilberg, C., Kiefer, J., Tillein, J., Pfennigsdorff, T., Hartmann, R., Stürzebecher, E. und Klinke, R. (1999). Electro-acoustic stimulation of the auditory system - New technology for severe hearing loss. *Journal of Otorhinolaryngology*, 61 (6), 334 - 340.
- Wagener, K., Kühnel, V. und Kollmeier, B. (1999a). Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests. *Zeitschrift für Audiologie*, 38 (1), 4 - 15.
- Wagener, K., Brand, T. und Kollmeier, B. (1999b). Entwicklung und Evaluation eines Satztests für die deutsche Sprache II: Optimierung des Oldenburger Satztests. *Zeitschrift für Audiologie*, 38 (2), 44 - 56.
- Wagener, K., Brand, T. und Kollmeier, B. (1999c). Entwicklung und Evaluation eines Satztests für die deutsche Sprache II: Evaluation des Oldenburger Satztests. *Zeitschrift für Audiologie*, 38 (3), 86 - 95.

- Wagener, K. und Brand, T. (2005). Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: influence of measurement procedure and masking parameters. *International Journal of Audiology*, 44 (3), 144 - 156.
- Watson, C.S., Kelly, W.J. und Wroton, H.W. (1976). Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty. *Journal of the Acoustical Society of America*, 60 (5), 1176 - 1186.
- Watson, C.S. (2005). Some comments on informational masking. *Acta Acustica united with Acustica*, 91, 502 - 512.
- Wegel, R.L. und Lane, C.E. (1924). The auditory masking of one sound by another and its probable relation to the dynamics of the inner ear. *Physical Review Letters*, 23, 266 - 285.
- Wendler, J., Seidner, W. und Eysholdt, U. (2005). *Lehrbuch der Phoniatrie und Pädaudiologie (4. Auflage)*. Stuttgart: Thieme Verlag KG.
- Wilson, B. S., Finley, C.C., Lawson, D.T., Wolford, R., Eddington, D.K. und Rabinowiz, W.M. (1991). Better speech recognition with cochlear implants. *Nature*, 352, 236 - 238.
- Wilson, B.S. und Dorman, M.F. (2008). Cochlear implants: Current designs and future possibilities. *Journal of Rehabilitation Research and Development*, 45 (5), 695 - 730.
- Wolf, A. (2004). *Sprachverstehen mit Cochlea-Implantat: EKP-Studien mit postlingual ertaubten erwachsenen CI-Trägern*. Dissertation, Universität Potsdam.
- Zeng, F.-G. (2011). Advances in auditory prostheses. In Zeng, F.-G., Popper, A.N. und Fay, R.R. (Hrsg.). *Auditory prostheses – New Horizons*. New York: Springer, 1 - 11.
- Zhang, T., Dorman, M.F. und Spahr, A.J. (2010). Information from the voice fundamental frequency (F0) region accounts for the majority of the benefit when acoustic stimulation is added to electric stimulation. *Ear and Hearing*, 31 (1), 63 - 69.
- Zierhofer, C.M. (1999). Electrical nerve stimulation based on channel specific sampling sequences. *Patent US application 60/150.773*.
- Zeng, F.-G. (2002). Temporal pitch in electric hearing. *Hearing Research*, 174, 101 - 106.

# Anhang A

## Verzeichnisse

### A.1 Abkürzungen

ACE	.....	Advanced Combination Encoder
ADC	.....	Analog-Digital Coder
CI	.....	Cochlea-Implantat
CIS	.....	Continuous Interleaved Sampling
CL	.....	Comfort Level (elektrische Behaglichkeitsschwelle)
dB	.....	Dezibel
dB HL	.....	Hearing Level (Hörschwelle) in Dezibel
dB SPL	.....	Sound Pressure Level (Schalldruck) in Dezibel
DAC	.....	Digital-Analog Coder
DBN	.....	Different-band-noise masker (Filterungsmethode)
DBS	.....	Different-band-sentence masker (Filterungsmethode)
EAS	.....	Elektro-Akustische Stimulation
EM	.....	Energetic Masking
FSP	.....	Fine Structure Processing
F <sub>0</sub>	.....	Grundfrequenz
F <sub>n</sub>	.....	Formantfrequenzen
HG	.....	Hörgerät
IM	.....	Informational Masking
ILD	.....	Interaural Level Difference (interaurale Pegeldifferenz)
ITD	.....	interaural time difference (interaurale Zeitdifferenz)
KS	.....	Kopfschatten
NH	.....	Normalhörend(e)
OISa	.....	Oldenburger Satztest

## Anhang A

---

pps	.....	pulses per second
SBN	.....	same-band noise masker (Filterungsmethode)
SNR	.....	Signal-to-Noise Ratio (Signal-Rauschabstand)
SPEAK	.....	Spectral Peak
SPL	.....	Sound pressure level (Schalldruckpegel)
SQ	.....	Squelch
SU	.....	Summation
THL	.....	Threshold Level (elektrische Hörschwelle)
TMR	.....	Target-to-Masker Ratio (Ziel-Maskierer Abstand)
VTL	.....	Vokaltraktlänge
WHO	.....	World-Health Organisation

### A.2 Glossar

**afferent** Bezeichnung für Nervenfasern, die eine Erregung der Peripherie an das zentrale Nervensystem weiterleiten; Gegensatz → efferent

**Aufblähkurve** Darstellung der Hörleistung eines Patienten im Tonaudiogramm, der mit einem Hörgerät versorgt ist. Eine Aufblähkurve wird im Freifeld gemessen.

**Bandpass** Ein Filter, der nur bestimmte Frequenzbereiche von akustischen Signalen überträgt. Frequenzen, die ober- oder unterhalb eines bestimmten Durchlassbereiches liegen, werden abgedämpft.

**Behaglichkeitsschwelle (CI)** Stellt die Schwelle bei einer Einzelkanalstimulation dar, bei der die Stromladung einen subjektiv lauten, aber noch angenehmen Reiz hervorruft.

**bilateral** Hören mit zwei Hörhilfen der gleichen Art; zentrale Integration liegt nicht unbedingt vor

**bimodal** Hören mit zwei Hörhilfen, die von unterschiedlicher Art sind; zentrale Integration liegt nicht unbedingt vor

**binaural** Hören mit beiden Ohren; die Signale die an jedem Ohr ankommen, werden zentralnervös verglichen und zu einem Objekt der Wahrnehmung integriert.

**biphasisch** aus zwei gegenläufigen Phasen bestehend, z.B. eine positive und eine negative Phase

**bottom-up** aufsteigender Verarbeitungsprozess beim Hören; akustische Merkmale bekommen durch den gleichzeitig ablaufenden → top-down Prozess Bedeutung

**Cochlea-Implantat** Innenohrprothese, die durch elektrische Stimulation den Hörnerv reizt und dadurch einen Höreindruck hervorruft.

**Cochleostomie** Zugang für die Insertion eines Cochleaelektrodenträgers. Hierbei wird die Hörschnecke an einer Stelle künstlich geöffnet.

**Cocktail-Party Effekt** beschreibt die Fähigkeit, dass aus mehreren simultanen Schallsignalen ein Signal extrahiert werden kann.

**closed set** Eine Methode bei Sprachtests, bei der der Patient Antwortalternativen vorgegeben bekommt, aus denen er die richtige Antwort wählen muss. Im Gegensatz hierzu steht die Methode „open set“, bei der der Patient aus seinem Gesamtwortschatz antworten muss ohne Vorgabe von Alternativen.

**diotisch** beide Ohren werden identisch beschallt

**Dynamikbereich** In einem Audiosystem der Bereich zwischen minimalem und maximalem Pegel, in dem sich ein Tonsignal „nutzbringend“ bewegen kann. Angaben in dB.

**efferent** Bezeichnung für Nervenfasern, die Impulse vom zentralen Nervensystem an die Peripherie weiterleiten; Gegensatz → afferent

**Einhüllende** Allgemeiner Umriss eines Klangsignals, der Veränderungen in der Lautstärke im zeitlichen Verlauf wiedergibt (<http://www.medel.com/de/finehearing/>)



**Energetic Masking** Störwirkung, die durch die spektrale Energie eines gleichzeitigen Schallsignals entsteht, durch das ein Zielsignal unhörbar wird.

**Feinstruktur** enthält schnell wechselnde Tonhöhendetails (bezogen auf Sprache).  
(<http://www.medel.com/de/finehearing/>)

**Forced-choice Methode** Vorgehensweise bei Tests, bei der Antwortalternativen vorgegeben werden.

**Formant(frequenz)** Eine Konzentration der Schallenergiespitzen in Spektrogrammen, die für Vokale und stimmhafte Laute charakteristisch ist.

**Freiburger Einsilber Test** Sprachverständlichkeitstest bestehend aus einsilbigen Nomen, bei dem der Patient auf seinen Gesamtwortschatz zurückgreifen muss.

**Freifeld** Schallsituation ohne Auftreten von Reflexionen.

**Grundfrequenz** Niedrigste Frequenzkomponente in einer komplexen Schallwelle; wichtig für die Bestimmung der Tonhöhe eines Lautes.

**Hörgerät** Hörhilfe, die durch Verstärkung des Schalldruckpegels versucht eine Hörstörung auszugleichen → s. Luftleitungshörgerät; Knochenleitungshörgerät.

**Hörschwelle** Schalldruckpegel, bei dem für eine bestimmte gemessene Frequenz im Gehör eben gerade eine Empfindung ausgelöst wird.

**Informational Masking** Störwirkung bei überlagerten Schallsignalen, die durch die Information des Maskierersignals zustande kommt. Diese Information ist irrelevant für den Hörer und stört das Verstehen des eigentlichen Zielsignals.

**Ipsilateral** auf der gleichen Seite gelegen; Gegensatz → kontralateral

**Kontralateral** auf der gegenüberliegenden Seite (Körperhälfte) gelegen; Gegensatz → ipsilateral

**Knochenleitungshörgerät** Hörgerät, das durch Vibration des Schädelknochens direkt eine Wanderwelle in der Cochlea hervorruft.

**Kopfschatten** physikalischer Effekt, der entsteht, wenn Ziel- und Störsignal räumlich getrennt sind. Ein Ohr liegt durch den Kopf im Schallschatten, so dass ein Störsignal abgedämpft wird.

**Lateral** seitlich von der Körpermitte; hier: vorrangige Zuständigkeit einer Hemisphäre des Gehirns.

**Logatom** meist einsilbiges Kunstwort zur Testung des Silbenverstehens, das die phonotaktischen Regeln der jeweiligen Sprache befolgt, aber kein Wort in der zu testenden Sprache darstellt.

**Luftleitungshörgerät** Ein Hörgerät, bei dem ein Hörer (Lautsprecher) Signale direkt an das Ohr, meist über ein Ohrpassstück, in den Gehörgang sendet.

**monaural** einohrig

## Anhang A

---

**Unbehaglichkeitsschwelle** Schalldruckpegel, bei dem im Gehör schmerzhaft empfindungen hervorgerufen werden

**Signal-Rausch Abstand (SNR)** Verhältnis zwischen Nutz- und Störsignal gemessen in dB.

**Target-to-masker ratio (TMR)** meist ein Begriff für sprachliche Signale, um den → Signal-Rausch Abstand eines Ziels und eines Maskierers zu beschreiben

**prä-, peri-, postlingual** Zeitpunkt des Eintritts einer Hörschädigung bezogen auf den Spracherwerb

**progredient** fortschreitend; in Bezug auf Hörstörungen bedeutet dies ein zunehmend schlechteres Hören mit der Zeit

**Prosodie** rhythmische und melodische Eigenschaften der Sprache, die über ein lautliches Segment hinausgehen

**Rekrutment** Lautheitsausgleich; bei Hörstörungen mit Ausfall der äußeren Haarzellen werden trotz herabgesetzter Hörschwelle, starke Schallsignale schnell als laut wahrgenommen. Es kommt zu einem schnelleren Lautheitsanstieg im Hörbereich im Vergleich zu Normalhörenden.

**Tonaudiometrie** misst die subjektive Hörschwelle von Personen in bestimmten Frequenzbereichen

**Tonhöhe** Hörempfindung, die eng mit der Frequenz eines Tones verbunden ist. Je höher die Frequenz ist, desto höher die Tonhöhenempfindung.

**Signal-Rausch Abstand** ergibt sich beim Sprachverstehen im Störgeräusch und bezeichnet den Abstand zwischen Ziel- und Maskierersignal in dB, bei dem 50% des Zielsignals verstanden wurden; Positive Werte zeigen an, dass der Pegel des Zielsignals höher war als der Pegel des Maskierersignals, negative Werte zeigen an, dass der Pegel des Zielsignals niedriger war als der des Maskierersignals.

**Squelch** Binaurale Störgeräuschunterdrückung, die durch die vergleichende Verarbeitung beider Ohren und somit beider aufsteigender Hörbahnen entsteht.

**Sprachaudiometrie** Untersucht die Verständlichkeit Sprache zu verstehen

**Summation** Zentrale Integration der Information von mehreren Schallsignalen, die an beiden Ohren ankommen. Die Schallsignale sind räumlich nicht getrennt.

**Tinnitus** subjektives oder objektives Ohrgeräusch

**Top-down** absteigender Verarbeitungsprozess beim Hören; wird vom Wissen und von den Erwartungen des Hörenden gesteuert. Gegenteil: → bottom-up

## Anhang B

### Methodenbezogene Information

#### B.1 Stimuligenerierung für den Vortest

Original Mann	F0	Halbtöne	VTL
Super-male	0,74	- 5,2	1,08
Male	1,00	0,0	1,00
Quarter-female	1,17	2,7	0,96
Half-female	1,35	5,2	0,92
Almost-female	1,53	7,4	0,88
Female	1,70	9,2	0,84
Super-female	2,05	12,4	0,76

Tabelle B.1: Originalkategorien für die Änderungen in F0 und VTL zur Generierung einer weiblichen Stimme aus einer männlichen Stimme (nach Darwin et al., 2003)

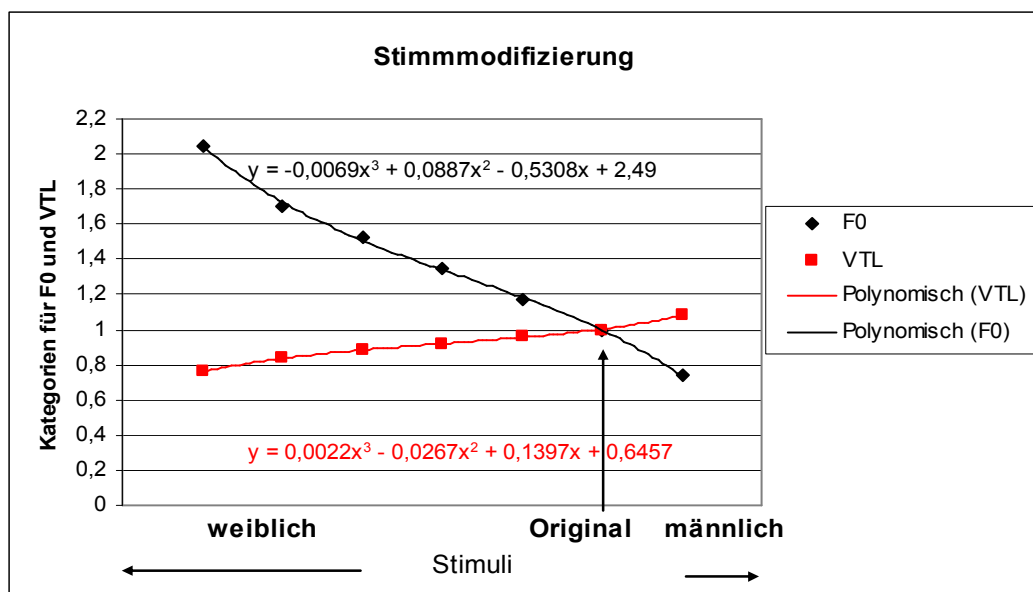


Abbildung B.1: Stimmmodifizierung durch Änderung von F0 und VTL. Graphische Darstellung der Werte aus Tabelle B.1. Die Werte wurden mithilfe zweier polynomischer Formeln (F0 und VTL) angenähert.

## Anhang B

---

Die Originalwerte von Darwin et al. (2003) konnten mithilfe einer polynomischen Formel 3. Grades angenähert werden. Die polynomische Formel B.1 wurde für die Interpolation der F0-Werte herangezogen.

$$y = 0,0069x^3 + 0,0887x^2 + 0,5308x + 2,49 \quad (\text{B.1})$$

Mithilfe der polynomischen Formel B.2 konnten die Werte für den Parameter VTL interpoliert werden

$$y = 0,0022x^3 + 0,0267x^2 + 0,1397x + 0,6457 \quad (\text{B.2})$$

Die nachfolgende Tabelle B.2 zeigt die interpolierten Werte und die dazugehörigen F0- und F<sub>n</sub>-Werte (F1 – F3). Alle Stimuli wurden für Experiment 1 (Sprecherunterscheidung) im Vortest (Kapitel 7) genutzt. Bei den eingerahmten Werten handelt es sich um die reduzierte Stimulanzahl, die in Experiment 2 (Sprechererkennung) des Vortests verwendet wurde.

F0	VTL	F0 [Hz]	F1 [Hz]	F2 [Hz]	F3 [Hz]
1,87	0,8	294	968	2194	3440
1,7	0,84	267	936	2120	3326
1,61	0,86	253	920	2084	3268
1,53	0,88	240	904	2047	3211
1,42	0,9	223	887	2011	3154
1,35	0,92	212	872	1974	3096
1,27	0,93	199	863	1956	3068
1,17	0,96	176	839	1901	2982
1,11	0,97	174	831	1883	2953
<b>1</b>	<b>1</b>	<b>157</b>	<b>807</b>	<b>1828</b>	<b>2867</b>
0,91	1,03	143	783	1773	2781
0,85	1,04	133	775	1755	2752
0,77	1,07	121	751	1700	2666
0,74	1,08	116	742	1682	2638
0,70	1,10	110	726	1645	2580
0,66	1,12	104	710	1609	2523
0,64	1,14	100	694	1572	2466
0,59	1,16	93	678	1536	2408
0,55	1,20	86	646	1462	2294

Tabelle B.2: Interpolierte Werte für das Sprecherunterscheidungsexperiment (Kapitel 7).  
Originalwerte: F0: 157 Hz; F1: 807 Hz; F2: 1828 Hz; F3: 2867 Hz

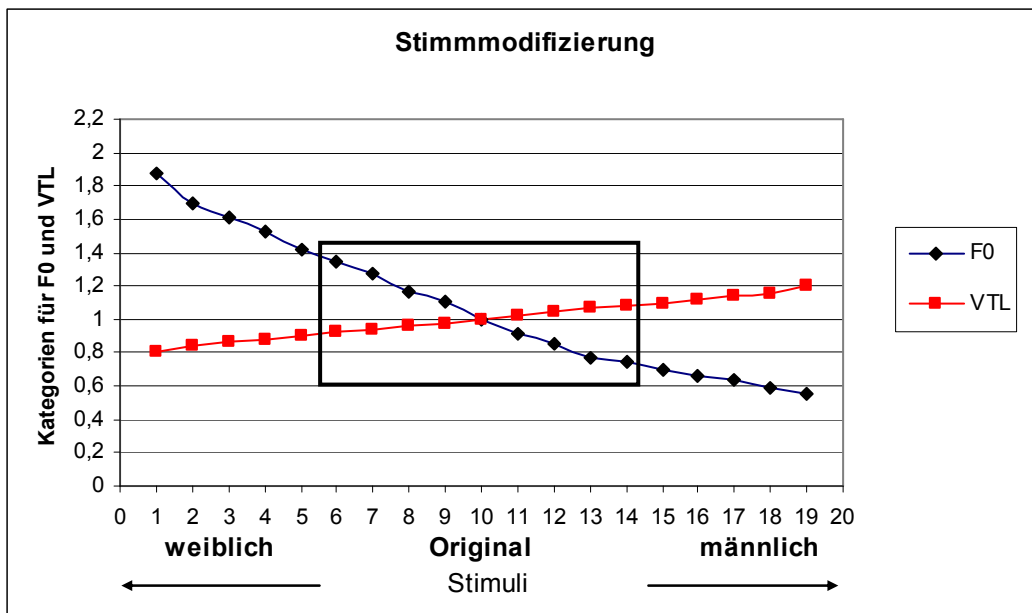


Abbildung B.2: Stimmmodifizierung durch Änderung von F0 und VTL. Graphische Darstellung der Werte aus Tabelle B.2 für Experiment 1 und (eingrahmt die Werte für Experiment 2 des Vortests (Kapitel 7)

### B.2 Skript zur Änderung von F0 und / oder VTL (nach Darwin et al., 2003)

```
# If the quality is bad try changing the minimum pitch or maximum pitch
# the script changes pitch and duration and then resamples to bring them
# back to original values.
```

```
form Change vocal tract size
  sentence Directory_with_waveform_files cjdk:disk:test:
  sentence Directory_for_output_files cjdk:disk:testout:
  choice Output_format: 1
    button AIFF
    button AIFC
    button WAV
  positive Lengthening_factor 1.0
  sentence Fo_expression self*1.0
  comment Analysis parameters
  positive minimum_Fo 75
  positive maximum_Fo 300
  boolean Play_after_synthesis 0
endform
```

```
vtfactor = 'Lengthening_factor'
directory$ = directory_with_waveform_files$
directoryo$ = directory_for_output_files$
fomin = 'minimum_Fo'
fomax = 'maximum_Fo'
```

```
Create Strings as file list... list 'directory$'*
numberOfFiles = Get number of strings
for ifile to numberOfFiles
  select Strings list
  sound$ = Get string... ifile
  call vtchange
endfor
select Strings list
Remove
```

```
procedure vtchange
#
Read from file... 'directory$''sound$'
#initial analysis locating pitch pulses etc
#strip off extension (defined as everything after last .
if rindex(sound$, ".") > 0
  sound$ = left$(sound$, rindex(sound$, ".")-1)
endif
select Sound 'sound$'
durn = Get duration

#create Pitch & Manipulation objects
select Sound 'sound$'
To Pitch... 0.01 fomin fomax
plus Sound 'sound$'
To Manipulation
select Pitch 'sound$'
#apply the appropriate transformation to the Pitch object
Formula... 'fo_expression$'*'vtfactor'
```

## Anhang B

---

```
#turn it into a PitchTier and place it into the Analysis object
select Pitch 'sound$'
Down to PitchTier
select Manipulation 'sound$'
plus PitchTier 'sound$'
Replace pitch tier

Create DurationTier... newdurn 0 'durn'
Add point... 0 1/'vtfactor'
select Manipulation 'sound$'
plus DurationTier newdurn
Replace duration tier

#resynthesise with new pitch and duration contour
select Manipulation 'sound$'
Get resynthesis (PSOLA)

#resample and force return of sample rate to original
fsamp = Get sample rate
Resample... fsamp*vtfactor 50
Override sample rate... fsamp
if play_after_synthesis = 1
    Play
endif

#tidy up decimal representation in name
vtf$ = fixed$(vtfactor,2)

# remove self* string if it is there
pos = index (fo_expression$,"self*")
if pos > 0
    eleng = length (fo_expression$)
    foe$ = left$ (fo_expression$, pos -1) + right$(fo_expression$,eleng-pos-4)
else
    foe$ = fo_expression$
endif

if output_format = 1
    Write to AIFF file... 'directoryo$''sound$'. 'vtf$'. 'foe$'.aiff
elseif output_format = 2
    Write to AIFC file... 'directoryo$''sound$'. 'vtf$'. 'foe$'.aifc
else
    Write to WAV file... 'directoryo$''sound$'. 'vtf$'. 'foe$'.wav
endif

select all
minus Strings list
Remove
endproc
```

### B.3 Basiswortschatz des Oldenburger Satztests

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

Tabelle B.3: Wortschatz und Konstruktion der Basisliste des Oldenburger Satztests - modifiziert (nach Wagener et al., 1999: 9).



## B.4 Modifizierte Listen des Oldenburger Satztest für IM

### Testliste 1

#### Zielsatz

- 1) Stefan hat zwei alte Messer.
- 2) Stefan verleiht elf grüne Autos.
- 3) Stefan nahm zwölf nasse Autos.
- 4) Stefan sieht vier grüne Tassen.
- 5) Stefan gibt sieben rote Sessel.
- 6) Stefan schenkt elf schöne Autos.
- 7) Stefan nahm fünf große Autos.
- 8) Stefan gewann sieben teure Bilder.
- 9) Stefan hat zwei große Blumen.
- 10) Stefan verleiht neun teure Autos.
- 11) Stefan verleiht fünf teure Ringe.
- 12) Stefan malt elft weiße Dosen.
- 13) Stefan gewann achtzehn große Sessel.
- 14) Stefan nahm neun teure Steine.
- 15) Stefan sieht vier kleine Schuhe.

#### Maskierersatz

- Britta nahm zwölf nasse Sessel.  
Tanja malt achtzehn schwere Dosen.  
Tanja kauft fünf rote Messer.  
Tanja hat zwei alte Messer.  
Nina bekommt drei kleine Bilder.  
Kerstin nahm acht kleine Dosen.  
Ulrich malt elf weiße Dosen.  
Peter kauft achtzehn nasse Schuhe.  
Ulrich schenkt sieben alte Messer.  
Kerstin bekommt zwei rote Tassen.  
Kerstin schenkt drei schwere Blumen.  
Doris gewann achtzehn kleine Bilder.  
Kerstin verleiht neun schwere Ringe.  
Thomas schenkt elf schöne Autos.  
Britta hat acht weiße Bilder.

### Testliste 2

#### Zielsatz

- 1) Stefan sieht achtzehn weiße Bilder.
- 2) Stefan schenkt fünf alte Ringe.
- 3) Stefan kauft drei große Ringe.
- 4) Stefan verleiht sieben große Ringe.
- 5) Stefan gewann elf schöne Dosen.
- 6) Stefan verleiht neun schwere Ringe.
- 7) Stefan sieht achtzehn grüne Sessel.
- 8) Stefan gibt vier weiße Tassen.
- 9) Stefan gewann elf schöne Blumen.
- 10) Stefan schenkt acht kleine Steine.
- 11) Stefan schenkt elf weiße Dosen.
- 12) Stefan schenkt drei schwere Blumen.
- 13) Stefan sieht zwei nasse Dosen.

#### Maskierersatz

- Wolfgang verleiht drei alte Sessel.  
Peter bekommt vier grüne Messer.  
Doris malt acht teure Dosen.  
Doris gewann vier schöne Blumen.  
Tanja gibt vier schwere Steine.  
Kerstin sieht achtzehn grüne Sessel.  
Wolfgang verleiht sieben große Ringe.  
Doris hat zwölf große Blumen.  
Britta gibt sieben rote Sessel.  
Thomas gewann elf schöne Dosen.  
Peter verleiht vier kleine Bilder.  
Wolfgang sieht achtzehn schöne Ringe.  
Nina nahm neun teure Autos.

## **Anhang B**

---

- 14) Stefan malt achtzehn schwere Dosen.
- 15) Stefan gibt fünf alte Steine.

- Wolfgang kauft elf rote Steine.  
Thomas hat zwölf große Autos.

### **Testliste 3**

#### Zielsatz

- 1) Stefan malt vier schwere Dosen.
- 2) Stefan gewann zwölf grüne Blumen.
- 3) Stefan kauft neun schwere Tassen.
- 4) Stefan hat zwölf große Autos.
- 5) Stefan gewann drei schwere Steine.
- 6) Stefan gibt zwei schöne Messer.
- 7) Stefan hat zwei alte Messer.
- 8) Stefan schenkt zwölf schwere Tassen.
- 9) Stefan gibt zwei weiße Autos.
  
- 10) Stefan bekommt vier grüne Messer.
- 11) Stefan malt drei große Ringe.
- 12) Stefan kauft fünf rote Messer.
- 13) Stefan verleiht drei alte Sessel.
- 14) Stefan kauft acht nasse Ringe.
- 15) Stefan sieht vier nasse Schuhe.

#### Maskierersatz

- Britta verleiht elf alte Bilder.  
Thomas kauft neun schwere Tassen.  
Ulrich verleiht sieben große Ringe.  
Tanja gibt drei rote Tassen.  
Thomas bekommt sieben rote Schuhe.  
Thomas nahm zwölf nasse Sessel.  
Nina schenkt drei teure Schuhe.  
Kerstin verleiht elf grüne Autos.  
Thomas bekommt achtzehn grüne Schuhe.  
Wolfgang kauft sieben schöne Schuhe.  
Nina gibt zwei weiße Autos.  
Thomas gewann achtzehn große Sessel.  
Tanja malt acht weiße Dosen.  
Tanja sieht drei rote Blumen.  
Thomas hat zwei große Blumen.

### **Testliste 4**

#### Zielsatz

- 1) Stefan nahm sieben schöne Tassen.
- 2) Stefan gibt fünf grüne Steine.
- 3) Stefan hat fünf grüne Bilder.
- 4) Stefan gibt elf rote Blumen.
- 5) Stefan bekommt zwei rote Tassen.
- 6) Stefan kauft elf rote Steine.
- 7) Stefan malt drei alte Messer.
- 8) Stefan gewann neun schöne Bilder.
- 9) Stefan verleiht vier kleine Bilder.
- 10) Stefan malt acht teure Dosen.
- 11) Stefan gibt vier schwere Steine.

#### Maskierersatz

- Doris bekommt acht nasse Steine.  
Doris verleiht zwölf kleine Bilder.  
Nina kauft zwei alte Ringe.  
Doris kauft drei große Ringe.  
Kerstin kauft acht nasse Ringe.  
Nina verleiht fünf teure Ringe.  
Tanja schenkt acht kleine Steine.  
Tanja schenkt zwei alte Tassen.  
Thomas bekommt fünf teure Sessel.  
Ulrich gibt elf rote Blumen.  
Ulrich gewann zwölf rote Messer.

## **Anhang B**

---

- 12) Stefan sieht zwei große Tassen.
- 13) Stefan bekommt sieben rote Schuhe.
- 14) Stefan bekommt elf schöne Messer.
- 15) Stefan verleiht zwölf kleine Bilder.

- Britta gewann drei schwere Steine.
- Kerstin nahm fünf große Autos.
- Nina sieht vier grüne Tassen.
- Ulrich bekommt achtzehn grüne Schuhe.

## B.5 Demographie der Probanden

### B.5.1 Vortest + Experiment 1

Probanden ID	Alter in Jahren	Ätiologie	Hersteller, Implantat	Sprachprozessor	Kodierungsstrategie	Höralter mit CI in Jahren
CI 01	39	Penicillin-unverträglichkeit, progredient	Advanced Bionics, HiRes 90k	Auria	HiRes-P	2;0
CI 02	39	unklar, progredient	Advanced Bionics, HiRes 90k	Auria	HiRes-S	0;10
CI 03	51	Infektion, progredient	Advanced Bionics, HiRes 90k	Auria	HiRes-P	2;5
CI 04	48	unklar, progredient	Cochlear, Nucleus CI24M	ESprit 3G	ACE	3;4
CI 05	43	unklar, progredient	Medel, Sonata	Tempo+	CIS	1;5
CI 06	69	unklar, progredient	Cochlear, CI24RE	Freedom	ACE	0;11

Tabelle B.5.1: Allgemeine Informationen zu den 6 unilateral versorgten CI-Probanden, die an dem Vortest und an Experiment 1 teilnahmen.

B.5.2 Experiment 2

Probanden ID	Alter in Jahren	Ätiologie	Hersteller, Implantat	Prozessor	Kodierungsstrategie	Höralter CI in Jahren	HG Typ	Höralter HG in Jahren
CI 01	69	Unklar	Cochlear, CI24RE	Freedom	ACE	1;4	Siemens, Triano SP	27;0
CI 02	64	Medikation nach Tuberkulose	Advanced Bionics, CII	PSP	HiRes-S	6;3	Siemens, Auriculina 224	30;0
CI 03	18	unklar	Med-EI, Pulsar	Opus II	FSP	1;2	Phonak, Eleva 411	17;0
CI 04	44	genetisch	Med-EI, Sonata	Opus II	FSP	0;3	Phonak, Savia 311	26;0
CI 05	67	genetisch	Med-EI, Pulsar	Opus II	FSP	0;5	Phonak, Valeo 311 AZ	18;0
CI 06	72	Trauma	Cochlear, CI24RE	Freedom	ACE	2;0	Siemens, Prisma H768	11;0
CI 07	64	unklar	Cochlear, CI24RE	Freedom	ACE	1;0	Siemens, Infiniti Pro SP	18;0
CI 08	23	unklar	Advanced Bionics, HiRes 90k	Auria	HiRes-S	0;10	Phonak, Naidalll UP	23;0
CI 09	59	Morbus Meniere	Cochlear, CI24RE	Freedom	ACE	1;2	Phonak, Naidalll UP	15;0
CI 10	54	unklar	Med-EI, Sonata	Opus II	FSP	0;4	Phonak, Naidalll UP	18;0
CI 11	63	Meningitis	Med-EI, Sonata	Opus II	FSP	0;8	Phonak, Perseo Forte	47;0
CI 12	71	unklar	Cochlear, CI512	CP810	ACE	0;5	Phonak, Supero	30;0

Tabelle B.5.2: Allgemeine Informationen zu den 12 bimodal versorgten CI-Probanden, die an Experiment 2 teilnahmen.

**B.5.3 Experiment 3**

Probanden ID	Alter in Jahren	Ätiologie	Hersteller, Implantat	Prozessor	Kodierungsstrategie	Höralter CI in Jahren	Hörgerät	Höralter HG in Jahren
CI 01	70	unklar	Cochlear, CI24RE	Freedom	ACE	2;4	Siemens, Triano SP	28;0
CI 02	19	unklar	Med-El, Pulsar	Opus II	FSP	2;2	Phonak, Eleva 411	17;0
CI 03	39	Röteln	Cochlear, CI24RE	Freedom	ACE	2;4	Phonak, Naida	38;0
CI 04	69	unklar	Cochlear, CI24RE	Freedom	ACE	0;10	Siemens, Infiniti Pro SP	18;0
CI 05	59	Morbus Meniere	Cochlear CI24RE	Freedom	ACE	0;10	Phonak, Naidalll UP	15;0
CI 06	63	Meningitis	Med-El, Sonata	Opus II	FSP	1;8	Phonak, Naida V UP	40;0
CI 07	27	unklar	Cochlear, CI512	CP810	ACE	1;4	Siemens, Cielo SP	21;0
CI 08	72	unklar	Cochlear CI512	CP810	ACE	1;8	Phonak, Supero 413	41;0
CI 09	46	unklar	Med-El, Concerto	Opus II	FSP	0;11	Oticon, Epoq P	9;0
CI 10	55	unklar	Cochlear, CI512	CP810	ACE	1;7	Phonak, Savia 311	29;0

Tabelle B.5.3: Allgemeine Informationen zu den 10 bimodal versorgten CI-Probanden, die an Experiment 3 teilnahmen.

## B.6 Audiogramme der bimodal versorgten Probanden

### B.6.1 Experiment 2

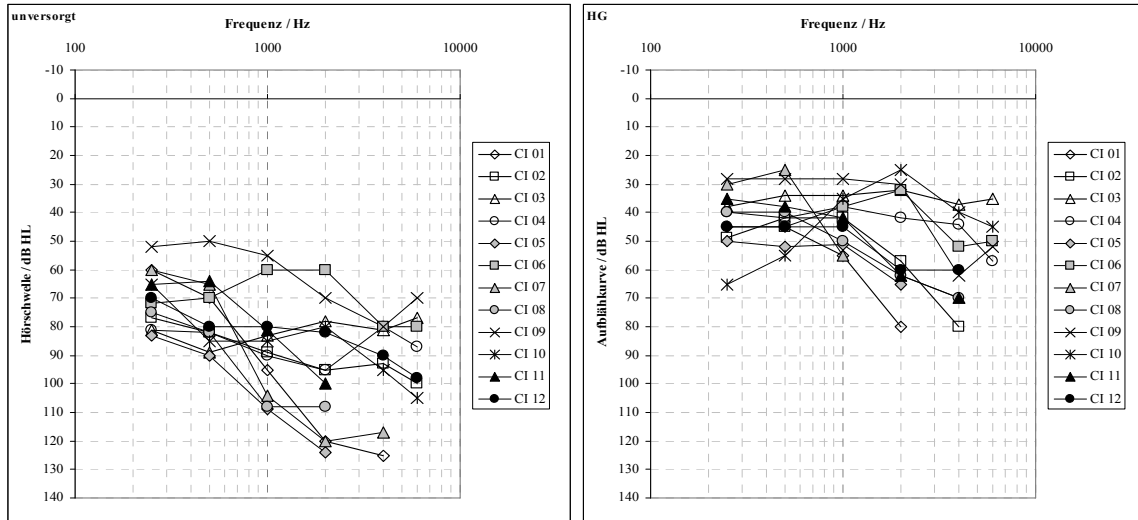


Abbildung B.6.1: Audiogramme der nicht-implantierten Seite der 12 bimodal versorgten CI-Probanden aus Experiment 2.

Links: unversorgte Hörschwelle; Rechts: Aufblähkurve mit Hörgerät

### B.6.2 Experiment 3

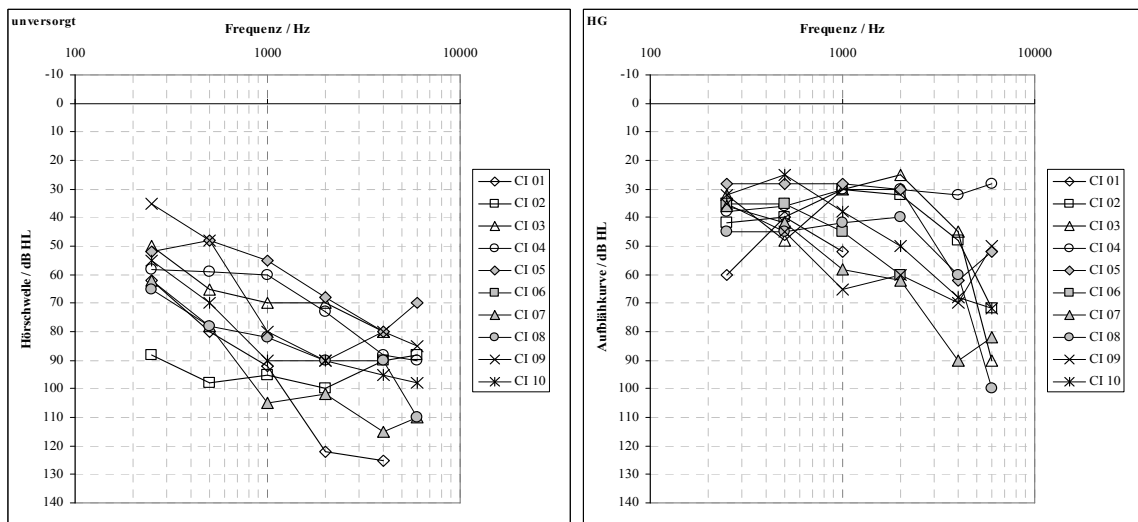


Abbildung B.6.2: Audiogramme der nicht-implantierten Seite der 10 bimodal versorgten CI-Probanden aus Experiment 3.

Links: unversorgte Hörschwelle; Rechts: Aufblähkurve mit Hörgerät.

## ANHANG C

### Statistische Auswertung

#### C.1 Experiment 1

##### C.1.1 Einfluss der Pegel- und Grundfrequenzunterschiede auf die Zielverständlichkeit

	Parameter	ANOVA	Ziel-Maskierer-Konditionen	Bonferroni
Zielverständlichkeit	<b>TMR</b>	<b>&lt;.001</b>	0 dB vs. 5 dB	<b>&lt;.001</b>
			0 dB vs. 10 dB	<b>&lt;.001</b>
			0 dB vs. 15 dB	<b>&lt;.001</b>
			0 dB vs. 20 dB	<b>&lt;.001</b>
			5 dB vs. 10 dB	<b>&lt;.001</b>
			5 dB vs. 15 dB	<b>&lt;.001</b>
			5 dB vs. 20 dB	<b>&lt;.001</b>
			10 dB vs. 15 dB	>.999
			10 dB vs. 20 dB	.096
			15 dB vs. 20 dB	>.999
	<b><math>\Delta F0</math></b>	.376		
	<b>TMR * <math>\Delta F0</math></b>	.995		

Tabelle C.1.1: p-Werte errechnet mithilfe der ANOVA für die Zielverständlichkeit der sechs unilateral versorgten CI-Probanden und die p-Werte der dazugehörigen Post-Hoc-Tests nach Bonferroni (für TMR bei  $\Delta F0 = 0$  Hz) (Signifikanzniveau:  $p = 0,05$ ).



**C.1.2 Einfluss der Pegel- und Grundfrequenzunterschiede auf den Anteil der A-Typ-Fehler**

	Parameter	ANOVA	Ziel-Maskierer-Konditionen	Bonferroni
A-Typ-Fehler	<b>TMR</b>	<b>&lt;.001</b>	0 dB vs. 5 dB	<b>&lt;.001</b>
			0 dB vs. 10 dB	<b>&lt;.001</b>
			0 dB vs. 15 dB	<b>&lt;.001</b>
			0 dB vs. 20 dB	<b>&lt;.001</b>
			5 dB vs. 10 dB	.504
			5 dB vs. 15 dB	>.999
			5 dB vs. 20 dB	<b>.036</b>
			10 dB vs. 15 dB	>.999
			10 dB vs. 20 dB	>.999
			15 dB vs. 20 dB	>.999
	<b><math>\Delta F0</math></b>	.073		
<b>TMR * <math>\Delta F0</math></b>	.270			

Tabelle C.1.2: p-Werte errechnet mithilfe der ANOVA für den Anteil der A-Typ-Fehler der sechs unilateral versorgten CI-Probanden und die p-Werte der dazugehörigen Post-Hoc-Tests nach Bonferroni (für TMR bei  $\Delta F0 = 0$  Hz) (Signifikanzniveau:  $p = 0,05$ ).

**C.1.3 Einfluss der Pegel- und Grundfrequenzunterschiede auf den Anteil der B-Typ-Fehler**

	Parameter	ANOVA	Ziel-Maskierer-Konditionen	Bonferroni
B-Typ-Fehler	<b>TMR</b>	<b>&lt;.001</b>	0 dB vs. 5 dB	<b>&lt;.001</b>
			0 dB vs. 10 dB	<b>&lt;.001</b>
			0 dB vs. 15 dB	<b>&lt;.001</b>
			0 dB vs. 20 dB	<b>&lt;.001</b>
			5 dB vs. 10 dB	<b>.012</b>
			5 dB vs. 15 dB	<b>&lt;.001</b>
			5 dB vs. 20 dB	<b>&lt;.001</b>
			10 dB vs. 15 dB	>.999
			10 dB vs. 20 dB	.572
			15 dB vs. 20 dB	>.999
	<b><math>\Delta F0</math></b>	.764		
<b>TMR * <math>\Delta F0</math></b>	.998			

Tabelle C.1.3: p-Werte errechnet mithilfe der ANOVA für den Anteil der B-Typ-Fehler der sechs unilateral versorgten CI-Probanden und die p-Werte der dazugehörigen Post-Hoc-Tests nach Bonferroni (für TMR bei  $\Delta F0 = 0$  Hz) (Signifikanzniveau:  $p = 0,05$ ).

## C.2 Experiment 2

### C.2.1 Einfluss der Hörhilfen auf die Zielverständlichkeit und die Fehlertypen

Ziel-Maskierer-Konditionen		ANOVA (Wilks Lambda)	CIHG - CI	CIHG - HG	CI - HG
0 Hz + 0%	Ziel	.075	>.999	.222	.174
	A-Typ	<b>.002</b>	<b>.018</b>	.183	.987
	B-Typ	<b>.044</b>	>.999	.123	.264
40 Hz + 0%	Ziel	<b>&lt;.001</b>	<b>.009</b>	<b>.027</b>	.534
	A-Typ	<b>.005</b>	.807	.174	<b>.030</b>
	B-Typ	<b>.002</b>	.369	<b>.024</b>	.129
80 Hz + 0%	Ziel	<b>&lt;.001</b>	<b>&lt;.001</b>	<b>.050</b>	.561
	A-Typ	<b>.013</b>	.249	.477	.060
	B-Typ	<b>.005</b>	<b>.036</b>	.057	.216
0 Hz + 8%	Ziel	.097	>.999	.324	.615
	A-Typ	.052	>.999	.114	.147
	B-Typ	<b>.027</b>	.900	.108	.237
0 Hz + 16%	Ziel	.162	>.999	.690	>.999
	A-Typ	<b>.004</b>	>.999	<b>.006</b>	<b>.027</b>
	B-Typ	<b>.019</b>	>.999	<b>.042</b>	.129
40 Hz + 8%	Ziel	<b>.003</b>	.120	.084	.873
	A-Typ	.100	>.999	.252	.282
	B-Typ	<b>.017</b>	.381	.102	.459
80 Hz + 16%	Ziel	<b>.002</b>	<b>.009</b>	.093	.897
	A-Typ	<b>.024</b>	>.999	<b>.045</b>	.057
	B-Typ	<b>.001</b>	.114	<b>.033</b>	.279

Tabelle C.2.1: p-Werte der ANOVA (Wilks Lambda) und der gepaarten Post-Hoc-Tests (nach Bonferroni) für die drei Hörhilfkonditionen HG, CI und CI + HG. Signifikante p-Werte sind hervorgehoben.

**C.2.2 Einfluss der Ziel-Maskierer-Konditionen auf die Zielverständlichkeit**

Zielverständlichkeit	Ziel-Maskierer-Konditionen	ANOVA (Wilks-Lambda)	Bonferroni (p-Werte)
unilateral CI	F0 0 Hz + 0% vs. 40 Hz + 0% 0 Hz + 0% vs. 80 Hz + 0% 40 Hz + 0% vs. 80 Hz + 0%	<b>.001</b>	>.999
			<b>.036</b>
			<b>.003</b>
	Fn 0 Hz + 0% vs. 0 Hz + 8% 0 Hz + 0% vs. 0 Hz + 16% 0 Hz + 8% vs. 0 Hz + 16%	<b>.003</b>	.117
			<b>.012</b>
			.738
bimodal CI + HG	F0 0 Hz + 0% vs. 40 Hz + 0% 0 Hz + 0% vs. 80 Hz + 0% 40 Hz + 0% vs. 80 Hz + 0%	<b>.001</b>	.102
			<b>.003</b>
			<b>.003</b>
	F0+Fn 0 Hz, + 0% vs. 40 Hz + 8% 0 Hz + 0% vs. 80 Hz + 16% 40 Hz + 8% vs. 80 Hz + 16%	<b>.016</b>	<b>.030</b>
			.114
			>.999

Tabelle C. 2.2: p-Werte der ANOVA (Wilks Lambda) und der gepaarten Post-Hoc-Tests (nach Bonferroni) für die Zielverständlichkeit bei den unterschiedlichen Ziel-Maskierer-Konditionen  $\Delta F_0$ ,  $\Delta F_n$  and  $\Delta F_0 + \Delta F_n$ . Signifikante p-Werte sind hervorgehoben.

**C.2.3 Einfluss der Ziel-Maskierer-Konditionen auf den Anteil der A-Typ-Fehler**

A-Typ-Fehler	Ziel-Maskierer-Konditionen	ANOVA (Wilks-Lambda)	Bonferroni (p-Werte)
unilateral CI	F0 0 Hz + 0% vs. 40 Hz + 0% 0 Hz + 0% vs. 80 Hz + 0% 40 Hz + 0% vs. 80 Hz + 0%	<b>.005</b>	>.999
			<b>.033</b>
			<b>.018</b>
	Fn 0 Hz + 0% vs. 0 Hz + 8% 0 Hz + 0% vs. 0 Hz + 16% 0 Hz + 8% vs. 0 Hz + 16%	<b>.028</b>	>.999
			.093
			.150
unilateral HG	F0 0 Hz + 0% vs. 40 Hz + 0% 0 Hz + 0% vs. 80 Hz + 0% 40 Hz + 0% vs. 80 Hz + 0%	<b>.004</b>	<b>.018</b>
			<b>.006</b>
			.097
	F0+Fn 0 Hz + 0% vs. 40 Hz + 8% 0 Hz + 0% vs. 80 Hz + 16% 40 Hz + 8% vs. 80 Hz + 16%	<b>.008</b>	.225
			<b>.015</b>
			.162
bimodal CI + HG	F0 0 Hz + 0% vs. 40 Hz + 0% 0 Hz + 0% vs. 80 Hz + 0% 40 Hz + 0% vs. 80 Hz + 0%	<b>&lt;.001</b>	<b>.009</b>
			<b>&lt;.001</b>
			<b>&lt;.001</b>
	F0+Fn 0 Hz + 0% vs. 40 Hz + 8% 0 Hz + 0% vs. 80 Hz + 16% 40 Hz + 8% vs. 80 Hz + 16%	<b>.029</b>	.186
			.060
			.600

Tabelle C. 2.3: p-Werte der ANOVA (Wilks Lambda) und der gepaarten Post-Hoc-Tests (nach Bonferroni) für die A-Typ-Fehler bei den unterschiedlichen Ziel-Maskierer-Konditionen  $\Delta F_0$ ,  $\Delta F_n$  and  $\Delta F_0 + \Delta F_n$ . Signifikante p-Werte sind hervorgehoben.

### C.3 Experiment 3

#### C.3.1 Einfluss der Effekte räumlichen Hörens auf die Zielverständlichkeit

Zielverständlichkeit					
Effekte		R 100 Hz	R 180 Hz	S 100 Hz	S 180 Hz
Kopfschatten	CI	<b>.005</b>	<b>.001</b>	<b>.000</b>	<b>.000</b>
	HG	<b>.009</b>	<b>.003</b>	<b>.001</b>	<b>.001</b>
Squelch	CI	<b>.012</b>	<b>.037</b>	.256	.959
	HG	.058	.429	<b>.026</b>	.904
Summation	CI	<b>.007</b>	<b>.011</b>	<b>.023</b>	.097
	HG	<b>.004</b>	<b>.021</b>	<b>.031</b>	<b>.002</b>

Tabelle C 3.1: p-Werte (T-Test) für die Zielverständlichkeit bei den Effekten KS, SQ und SU der 10 bimodal versorgten Probanden.

#### C.3.2 Einfluss der Maskierer, Hörhilfe und F0 auf die Zielverständlichkeit beim Kopfschatteneffekt

Kopfschatten	ANOVA (Wilks-Lambda)	Maskiererkondition	Hörkondition	Bonferroni p-Wert
Maskierer	<b>.008</b>	R100 vs. S100	CI	<b>.044</b>
		R180 vs. S180		.076
		R100 vs. S100	HG	.228
		R180 vs. S180		.700
Seite (Hörhilfe)	.098	R 100	CI vs. HG	>.999
		R 180		>.999
		S 100		.084
		S 180		.432
F0	.937	R 100 vs. R 180	CI	>.999
		S 100 vs. S 180		.400
		R 100 vs. R 180	HG	.444
		S 100 vs. S 180		>.999
Maskierer x Seite	.330			
Maskierer x F0	<b>.005</b>			
Seite x F0	.051			
Maskierer x Seite x F0	.630			

Tabelle C 3.2: p-Werte der ANOVA (Wilks Lambda) und der gepaarten Post-Hoc-Tests (nach Bonferroni) für die Zielverständlichkeit beim Kopfschatteneffekt. Einflussfaktoren: Maskierer, Hörhilfe, Grundfrequenz. Signifikante p-Werte sind hervorgehoben.

**C.3.3 Einfluss der Maskierer, Hörhilfe und F0 auf die Zielverständlichkeit beim Squelcheffekt**

Squelch	ANOVA	Maskiererkondition	Hörkondition	p-Wert
Maskierer	<b>&lt;.001</b>	R100 vs. S100	CI	<b>.016</b>
		R180 vs. S180		<b>.020</b>
		R100 vs. S100	HG	.088
		R180 vs. S180		>.999
Seite	.059	R 100	CI vs. HG	.412
		R 180		.636
		S 100		.504
		S 180		>.999
F0	.109	R 100 vs. R180	CI	>.999
		S 100 vs. S180		>.999
		R 100 vs. R180	HG	>.999
		S 100 vs. S180		<b>.040</b>
Maskierer x Seite	.428			
Maskierer x F0	<b>.002</b>			
Seite x F0	.200			
Maskierer x Seite x F0	.190			

Tabelle C 3.3: p-Werte der ANOVA (Wilks Lambda) und der gepaarten Post-Hoc-Tests (nach Bonferroni) für die Zielverständlichkeit beim Squelcheffekt. Einflussfaktoren: Maskierer, Hörhilfe, Grundfrequenz. Signifikante p-Werte sind hervorgehoben.

**C.3.4 Einfluss der Maskierer, Hörhilfe und F0 auf die Zielverständlichkeit beim Summationseffekt**

Summation	ANOVA	Maskiererkondition	Hörkondition	p-Wert
Maskierer	<b>.001</b>	R100 vs. S100	CI	.056
		R180 vs. S180		.084
		R100 vs. S100	HG	<b>.048</b>
		R180 vs. S180		.734
Seite	847	R 100	CI vs. HG	>.999
		R 180		>.999
		S 100		>.999
		S 180		>.999
F0	.729	R 100 vs. R 180	CI	.812
		S 100 vs. S180		>.999
		R 100 vs. R180	HG	>.999
		S 100 vs. S180		.236
Maskierer x Seite	.183			
Maskierer x F0	.085			
Seite x F0	.111			
Maskierer x Seite x F0	.111			

Tabelle C 3.4: p-Werte der ANOVA (Wilks Lambda) und der gepaarten Post-Hoc Tests (nach Bonferroni) für die Zielverständlichkeit beim Summationseffekt.

Einflussfaktoren: Maskierer, Hörhilfe, Grundfrequenz. Signifikante p-Werte sind hervorgehoben.

**C.3.5 Einfluss der Effekte räumlichen Hörens auf den Anteil der Fehler**

Fehleranalyse					
Effekte		A-Typ-Fehler		B-Typ-Fehler	
		S 100 Hz	S 180 Hz	S 100 Hz	S 180 Hz
Kopfschatten	CI	<b>.000</b>	<b>.000</b>	<b>.004</b>	<b>.001</b>
	HG	<b>.002</b>	<b>.004</b>	.088	<b>.006</b>
Squelch	CI	<b>.011</b>	<b>.015</b>	.233	.097
	HG	<b>.006</b>	<b>.046</b>	.696	.056
Summation	CI	<b>.024</b>	.108	<b>.019</b>	<b>.025</b>
	HG	<b>.002</b>	.430	<b>.003</b>	<b>.011</b>

Tabelle C.3.5: p-Werte (T-Test) für die Fehlertypen bei Verwendung der Sprachmaskierer für die Effekte KS, SQ und SU der 10 bimodalen Probanden.

**C.3.6 Einfluss der Maskierer, Hörhilfe und F0 auf den Anteil der Fehler beim Kopfschatteneffekt**

Kopfschatten	Fehlertyp	ANOVA	Maskierer-kondition	Hörkondition	p-Wert
Seite	A-Typ	.056	S 100	CI vs. HG	.444
			S 180		
F0		<b>.002</b>	S 100 vs. S 180	CI	<b>.004</b>
				HG	.072
Seite x F0		.787			
Seite	B-Typ	.182	S 100	CI vs. HG	.452
			S 180		
F0		<b>.030</b>	S 100 vs. S 180	CI	>.999
				HG	.104
Seite x F0		.384			

Tabelle C 3.6: p-Werte der ANOVA (Wilks Lambda) und der gepaarten Post-Hoc-Tests (nach Bonferroni) für die Fehlertypen beim Kopfschatteneffekt.

Einflussfaktoren: Maskierer, Hörhilfe, Grundfrequenz. Signifikante p-Werte sind hervorgehoben.

**C.3.7 Einfluss der Maskierer, Hörhilfe und F0 auf den Anteil der Fehler beim Squelcheffekt**

Squelch	Fehlertyp	ANOVA	Maskierer-kondition	Hörkondition	p-Wert
Seite	A-Typ	.870	S 100	CI vs. HG	>.999
			S 180		
F0		.060	S 100 vs. S 180	CI	>.999
				HG	.064
Seite x F0		.429			
Seite	B-Typ	.234	S 100	CI vs. HG	.560
			S 180		
F0		<b>.046</b>	S 100 vs. S 180	CI	>.999
				HG	.256
Seite x F0		.235			

Tabelle C 3.7: p-Werte der ANOVA (Wilks Lambda) und der gepaarten Post-Hoc-Tests (nach Bonferroni) für die Fehlertypen beim Squelcheffekt.

Einflussfaktoren: Maskierer, Hörhilfe, Grundfrequenz. Signifikante p-Werte sind hervorgehoben.

**C.3.8 Einfluss der Maskierer, Hörhilfe und F0 auf den Anteil der Fehler beim Summationseffekt**

Summation	Fehlertyp	ANOVA (Wilks Lambda)	Maskierer-kondition	Hörkondition	p-Wert
Seite	A-Typ	.343	S 100	CI vs. HG	>.999
			S 180		.104
F0		<b>.016</b>	S 100 vs. S 180	CI	.840
				HG	<b>.020</b>
Seite x F0		.190			
Seite	B-Typ	.912	S 100	CI vs. HG	>.999
			S 180		>.999
F0		.754	S 100 vs. S 180	CI	>.999
				HG	>.999
Seite x F0		.214			

Tabelle C 3.8: p-Werte der ANOVA (Wilks Lambda) und der gepaarten Post-Hoc-Tests (nach Bonferroni) für die Fehlertypen beim Squelcheffekt.

Einflussfaktoren: Maskierer, Hörhilfe, Grundfrequenz. Signifikante p-Werte sind hervorgehoben.