

Aus der Klinik und Poliklinik für Hals-, Nasen- und Ohren-  
krankheiten, plastische und ästhetische Operationen  
der Universität Würzburg

Direktor: Professor Dr. med. Rudolf Hagen

## **Evaluierung und Weiterentwicklung eines Prosodieverständnistests**

Inaugural-Dissertation  
zur Erlangung der Doktorwürde der  
Medizinischen Fakultät  
der  
Julius-Maximilians-Universität Würzburg  
vorgelegt von

Carmen Cordelia Domes

aus Marburg

Würzburg, Oktober 2009



Referent: Prof. Dr. med. Joachim Müller  
Korreferentin: Prof. Dr. rer. nat. Kathleen Wermke  
Dekan: Prof. Dr. med. Matthias Frosch

Tag der mündlichen Prüfung: 28. Juni 2010

Die Promovendin ist Zahnärztin





Meinen Eltern gewidmet, die mir mein Studium ermöglicht haben.



# Inhaltsverzeichnis

|          |                                                              |           |
|----------|--------------------------------------------------------------|-----------|
| <b>1</b> | <b>Einleitung</b>                                            | <b>1</b>  |
| 1.1      | Einführung . . . . .                                         | 1         |
| 1.1.1    | Prosodie – Begriff und Bedeutung . . . . .                   | 2         |
| 1.1.2    | Kochleaimplantate und ihre historische Entwicklung . . . . . | 4         |
| 1.2      | Aktueller Wissensstand . . . . .                             | 6         |
| 1.3      | Fragestellung und Zielsetzung . . . . .                      | 8         |
| <b>2</b> | <b>Versuchsplanung und Methodik</b>                          | <b>11</b> |
| 2.1      | Auswahl der Probanden . . . . .                              | 11        |
| 2.2      | Tonaudiogramm . . . . .                                      | 12        |
| 2.3      | Aufbau der Testsätze . . . . .                               | 13        |
| 2.3.1    | Verwendete Sprachdateien . . . . .                           | 13        |
| 2.3.2    | Störgeräusch . . . . .                                       | 13        |
| 2.3.3    | Erzeugung verrauschter Satzdateien . . . . .                 | 13        |
| 2.4      | Vorversuche . . . . .                                        | 16        |
| 2.4.1    | Erstellung und Kombination der Listen . . . . .              | 19        |
| 2.4.2    | Festlegung der absoluten Lautheitspegel . . . . .            | 21        |
| 2.4.3    | Auswertung der Vorversuche . . . . .                         | 21        |
| 2.5      | Hauptversuche . . . . .                                      | 24        |
| 2.5.1    | Kombination der Listen . . . . .                             | 26        |
| 2.5.2    | Festlegung der absoluten Lautheitspegel . . . . .            | 27        |
| 2.6      | Versuchsdurchführung . . . . .                               | 29        |
| 2.6.1    | Freifelddarbietung im schallgedämpften Raum . . . . .        | 29        |
| 2.6.2    | Probandenposition . . . . .                                  | 30        |
| 2.7      | Instruktion der Probanden . . . . .                          | 30        |
| 2.7.1    | Wortlaut der schriftlichen Probandeninstruktion . . . . .    | 31        |
| 2.8      | Versuchsablauf . . . . .                                     | 33        |
| 2.9      | Methodik der statistischen Auswertung . . . . .              | 33        |

|          |                                                                     |           |
|----------|---------------------------------------------------------------------|-----------|
| 2.9.1    | Darstellung der Gesamtdaten                                         | 34        |
| 2.9.2    | Urteilsverteilung                                                   | 36        |
| 2.9.3    | Verstehbarkeit der Sätze                                            | 37        |
| <b>3</b> | <b>Ergebnisse</b>                                                   | <b>39</b> |
| 3.1      | Darstellung der Gesamtdaten                                         | 39        |
| 3.2      | Urteilsverteilung                                                   | 42        |
| 3.2.1    | Urteilserkennung                                                    | 43        |
| 3.2.2    | Verwechslungen                                                      | 45        |
| 3.2.3    | Urteilshäufigkeit                                                   | 46        |
| 3.3      | Ratewahrscheinlichkeiten                                            | 46        |
| 3.4      | Verstehbarkeit der Sätze                                            | 50        |
| 3.4.1    | Erkennbarkeit der Kategorien                                        | 50        |
| 3.4.2    | Erkennbarkeit der Sprecher                                          | 51        |
| 3.5      | Empfindungen und Schwierigkeiten der Probanden während des Versuchs | 52        |
| <b>4</b> | <b>Diskussion</b>                                                   | <b>55</b> |
| 4.1      | Betrachtung der Gesamtdaten                                         | 55        |
| 4.2      | Betrachtung der Urteilsverteilung                                   | 57        |
| 4.2.1    | Urteilserkennung                                                    | 57        |
| 4.2.2    | Verwechslungen                                                      | 58        |
| 4.2.3    | Urteilshäufigkeit                                                   | 59        |
| 4.3      | Betrachtung der Ratewahrscheinlichkeiten                            | 60        |
| 4.4      | Betrachtung der Verstehbarkeit der Sätze                            | 64        |
| 4.4.1    | Erkennbarkeit der Kategorien                                        | 65        |
| 4.4.2    | Erkennbarkeit der Sprecher                                          | 66        |
| 4.5      | Fehlermöglichkeiten                                                 | 68        |
| <b>5</b> | <b>Ausblick</b>                                                     | <b>71</b> |
| <b>6</b> | <b>Zusammenfassung</b>                                              | <b>75</b> |
|          | <b>Anhang</b>                                                       | <b>77</b> |
| <b>A</b> | <b>Testsätze</b>                                                    | <b>79</b> |
| <b>B</b> | <b>Kontingenztabellen</b>                                           | <b>81</b> |

|                              |            |
|------------------------------|------------|
| <b>C Einzelsatzgraphiken</b> | <b>85</b>  |
| <b>D Satzvorschläge</b>      | <b>99</b>  |
| <b>Literaturverzeichnis</b>  | <b>101</b> |
| <b>Danksagung</b>            |            |



# 1 Einleitung

## 1.1 Einführung

Zur Behandlung der Schwerhörigkeit liegen zum gegenwärtigen Zeitpunkt verschiedene Methoden vor. In der Mehrzahl der Fälle wird je nach Art des Hörverlustes eine der folgenden Therapien angewendet. Zum einen kann nach Diagnosestellung einer geringgradigen Schwerhörigkeit bei intaktem Innenohr eine externe Hörhilfe zum Einsatz gebracht werden. Mit dieser wird die Schallwahrnehmung akustisch verstärkt. Bei Innenohrtaubheit und gesundem Hörnerv sowie zentralem Hörsystem werden Cochleaimplantate (CI) eingesetzt, die dem Patienten neue Hörleistungen ermöglichen sollen. Dabei ersetzt das CI als elektronische Hörprothese die ausgefallene Innenohrfunktion (vgl. [Boenninghaus und Lenarz, 2005](#)). Wird eine beidseitige Zerstörung des Hörnervs diagnostiziert, bleibt bei funktionstüchtigem zentralen Hörsystem nur die Hirnstammimplantation.

Mit diesen technischen Hörhilfen sind die meisten Patienten in der Lage, Satzstrukturen syntaktisch und grammatisch zu entschlüsseln. Weitgehend unberücksichtigt bleiben prosodische Aspekte des Sprachverstehens. Beispielsweise gibt es im deutschen Sprachgebrauch Satzkategorien, die auf ausschließlich grammatischer Basis nicht unterscheidbar sind. Für Normalhörende stellt dies kein Problem dar, da sie syntaktisch identische Sätze aufgrund der Wahrnehmung prosodischer Aspekte einer Frage, einem Ausruf oder einer Aussage zuordnen können. Auch die Erkennung von Stimmqualitäten spielt eine wichtige Rolle. Bei Telefongesprächen beispielsweise gibt der Inhalt des Gesagten keinerlei Aufschluss über die Person des Sprechers, so dass allein mit Hilfe der Tonhöhe als prosodischem Merkmal zwischen einem Mann, einer Frau und einem Kind unterschieden werden muss. Dies ist für Normalhörende in der Regel sehr einfach. Personen mit eingeschränkter Prosodiewahrnehmung weisen jedoch Restriktionen in der alltäglichen Kommunikation auf.

Nach [Meister et al. \(2007a\)](#) ist die Verarbeitung der der Prosodie zugrunde liegenden akustischen Parameter sowohl bei Patienten mit Hörstörungen als auch bei Hörhilfeträgern eingeschränkt. Unter den mit technischen Hilfsmitteln versorgten Personen sind besonders die CI-Träger hervorzuheben, da die speziellen Signalverar-

beitungsstrategien der Kochleaimplantate bezüglich der Wahrnehmung prosodischer Merkmale limitiert sind (Meister et al., 2007a). Zu den gleichen Ergebnissen kamen auch Wilson et al. (1991), Brill et al. (1997), Fishman et al. (1997), Garnham et al. (2002) und Kong et al. (2004). Um die Prosodiewahrnehmung bei Hörgeräte- und insbesondere bei CI-Trägern untersuchen zu können, sind adäquate Tests von großer Wichtigkeit. Bislang stehen diese im deutschsprachigen Raum nicht oder nur in sehr geringem Umfang zur Verfügung. Zudem muss besagter Test an normalhörenden Probanden als Referenzgruppe getestet werden. Hierdurch sollen zum einen Vergleichswerte Normalhörender ermittelt und zum anderen die Funktionstüchtigkeit und die Eigenschaften des Tests per se festgestellt werden. Verschiedene Signalverarbeitungsstrategien der Kochleaimplantate können bezüglich der Wahrnehmung prosodischer Merkmale miteinander verglichen und gegebenenfalls verbessert werden.

### 1.1.1 Prosodie – Begriff und Bedeutung

Zur besseren Verständlichkeit muss der Begriff „Prosodie“ zunächst erläutert werden. Aus phonetischer Sicht beschreibt die Prosodie – oftmals auch als „Suprasegmentalia“ bezeichnet – die Beiträge der Dauer, Grundfrequenz und der Amplitude zur Bedeutung der gesprochenen Sprache. Es werden sowohl paralinguistische Informationen wie z. B. zum emotionalen Zustand des Sprechers (froh, verärgert, nervös, etc.) als auch linguistische Informationen vermittelt. Linguistische Aspekte lassen beispielsweise den Perzipienten zwischen einer Aussage und einer Frage unterscheiden, auch wenn diese von der Satzkonstruktion identisch sind.

Möbius (1993) unterscheidet in seinem Modell der Prosodie die Suprasegmentalia Intonation, Lautheit und Quantität. Mit diesen linguistischen Parametern sind die akustischen Größen Grundfrequenz ( $f_0$ ), Intensität und Dauer eng korreliert. Andere prosodische Merkmale außer den Suprasegmentalia wie Sprechtempo, Rhythmus, Stimmqualität und Pausen werden in Möbius (1993) kaum diskutiert.

In der Literatur ist oft von Intonation anstelle der Prosodie die Rede. Dabei besteht Inozuka (2003) zufolge keine Übereinstimmung in der Definition. Einige Autoren verzichten sogar auf eine explizite Definition und setzen den Begriff als bekannt voraus. Da in der vorliegenden Arbeit insbesondere der Intonation als prosodischem Aspekt ein hoher Stellenwert zukommt, sollen an dieser Stelle einige Definitionen kurz erläutert werden.

Schon Jones (1960) beschäftigte sich mit Experimenten der Intonation, indem er



durch Stimmlippenschwingungen entstandene Tonhöhen notierte. Ihm zufolge „ist die Intonation die perzeptorische Tonhöhe selbst“.

Nach **Essen (1964)** ist die Intonation eine sprachliche Melodieregelung, die „nicht überlegt und nicht gewollt“ ist. Vielmehr ergibt sie sich „von selbst aus den seelischen Spannungsschwankungen, denen immer körperliche Spannungsänderungen parallel gehen“. Einige dieser Sprechmelodien sind „innerhalb einer Sprachgemeinschaft zum ‚Sprechmuster‘ erstarrt“ und werden als Intonation bezeichnet. Dabei ist die Intonation davon abhängig, was in einem Satz hervorgehoben werden soll. Neben Veränderungen der Dauer und des Atemdrucks ist die Variation der Tonhöhe das wichtigste Mittel zur Hervorhebung. Von Essen zufolge erstreckt sich die melodische Gestaltung nur über die Endphase einer Äußerung, in der zum Ausdruck gebracht wird, „welche Ausspruchskategorie gemeint ist“, beispielsweise die Ausspruchskategorie Aussage.

**Delattre et al. (1965)** setzen die Intonation mit dem Grundfrequenzverlauf gleich. Dazu muss angemerkt werden, dass andere akustische Parameter nicht erwähnt wurden, weil deren Anteile zum damaligen Forschungszeitpunkt noch nicht untersucht waren. Aus diesem Grund halten die Autoren die Gleichsetzung für gültig. Nach Delattre et al. weist die Intonation einen distinktiven und einen indistinktiven Aspekt auf. Dabei unterscheidet der distinktive Aspekt zwischen „Modi“, mit denen z. B. Frage und Befehl, aber auch Hervorhebungen von verschiedenen Wörtern innerhalb eines Satzes unterschieden werden. Durch die Worthervorhebungen, die von den Autoren als „Hauptkontinuation“, „Nebenkontinuation“ und „Termination“ bezeichnet werden, wird die Satzstruktur eindeutig.

Nach **Fery (1993)** setzt sich die Intonation aus verschiedenen Effekten zusammen. Tonhöhe, Phrasierungen, Akzente und Dauer bestehen nebeneinander in jeweils unterschiedlicher Ausprägung. Sie untersuchte ausschließlich Grundfrequenzverläufe, die sie als physikalische Realisation der Intonation bezeichnet und mit der Melodie einer Äußerung („intonational contour“) gleichsetzt.

Untersuchungen von **Möbius (1993)** beschreiben schließlich die Intonation als ein „suprasegmentelles Merkmalsystem“, dessen akustisches Korrelat die Sprachgrundfrequenz ist. Diese kann „meßtechnisch aus dem Sprachsignal extrahiert werden“.

### **Mikro- und Makroprosodie**

Prinzipiell kann zwischen Mikro- und Makroprosodie differenziert werden. Dabei handelt es sich bei der Mikroprosodie um nicht intendierte, segmentelle Einflüsse auf die Dauer, Amplitude und Grundfrequenz, während die Makroprosodie eine intendierte

## 1 Einleitung

und geplante Manipulation der genannten akustischen Parameter beschreibt. Mikroprosodische Einflüsse sind demzufolge vom Kontext vorhersagbar und phonetisch bedingt. Dagegen ist die Makroprosodie als kontextunabhängig und phonologisch zu betrachten. Während der Sprachwahrnehmung entfernt der Hörer unbewusst die mikroprosodischen Einflüsse.

Einer dieser mikroprosodischen Faktoren ist die vokalspezifische Grundfrequenz, die die Beeinflussung der spezifischen Grundfrequenz eines Vokals durch seine Qualität beschreibt (vgl. Inozuka, 2003). Lehiste und Peterson (1961) fanden in ihren Messungen heraus, dass bei Hochzungenvokalen die vokalspezifische Grundfrequenz höher ist als bei Tiefzungenvokalen. So ist die Grundfrequenz bei /i/ im Durchschnitt etwa um 20 Hz höher als bei /æ/. Nach Lehiste und Peterson (1961) können Konsonanten ebenfalls die Grundfrequenz folgender oder vorausgehender Vokale beeinflussen. Dabei erhöht ein stimmloser Konsonant die Grundfrequenzwerte des folgenden Vokals, während ein stimmhafter diese senkt. Geht ein /f/ beispielsweise einem /i/ voran, ist die Grundfrequenz um etwa 30 Hz höher als bei nach /v/ folgendem /i/. Diese beiden Phänomene werden von Möbius (1993) ebenfalls beschrieben und als „weitgehend sprachabhängig“ bezeichnet, da sie nicht der aktiven Kontrolle des Sprechers unterliegen.

Es stellt sich die Frage, inwieweit die Mikroprosodie perzeptorisch relevant ist und ob sie in intonationsbezogenen Untersuchungen Berücksichtigung finden sollte. Die Auffassungen in der Literatur hierüber sind zweigeteilt. Hart et al. (1990) und Adriaens (1991) z. B. sind der Meinung, dass die Mikroprosodie perzeptorisch irrelevant ist, während andere Autoren gegenteilige Ansichten vertreten (Inozuka, 2003).

### 1.1.2 Kochleaimplantate und ihre historische Entwicklung

Kochleaimplantate ersetzen die Funktion geschädigter Haarzellen des Innenohres. Dabei wandeln sie Schall in elektrische Signale um, die dann vom Hörnerv weiterverarbeitet werden können. Der Schall wird vom Mikrophon des am Ohr getragenen Sprachprozessors aufgenommen, der die Signale in ein spezielles elektrisches Pulsmuster kodiert und über ein Kabel zur Sendespule leitet. Diese wird gegenüber dem Implantat plaziert und übermittelt die kodierten Signale transkutan zum Implantat, welches das Pulsmuster entschlüsselt und über die Elektroden die Cochlea elektrisch stimuliert (vgl. Wilson und Dorman, 2008). Je nach Art des Implantats ist eine unterschiedliche Anzahl von Elektroden intra- oder extrakochleär positioniert, die durch elektrische Reize den Hörnerv stimulieren. Heute kommen nur noch intrakochleäre

Multielektrodenimplantate zum Einsatz.

In den letzten Jahren wurde die Qualität der Kochleaimplantate ständig verbessert. Dafür waren neben vielen verschiedenen Faktoren vor allem auch neue Signalverarbeitungs- und Stimulationsstrategien verantwortlich (vgl. [Mitterbacher, 2004](#)). Anfangs galt die Sprachverarbeitung mit der „compressed analogue“ (CA) Strategie als Standard für mehrkanalige Kochleaimplantate. Hierbei wurde das mit dem Mikrofon aufgenommene Signal in vier Frequenzbereiche gefiltert, deren Kanäle jeweils mit einer Elektrode verbunden waren. Die Informationen wurden so durch die relative Amplitude des Stimulus innerhalb der vier Kanäle und durch die zeitlichen Details der Wellenform kodiert. Als problematisch stellte sich heraus, dass nur ein Teil der sprachlichen Informationen vom Perzipienten auch tatsächlich wahrgenommen werden konnte. Weiterhin konnten durch die simultane Präsentation der Stimuli signifikante Interaktionen zwischen den Kanälen durch Vektorsummation der elektrischen Felder jeder Elektrode entstehen ([Wilson et al., 1991](#)).

Zur Verbesserung der Sprachwahrnehmung wurde Ende der achtziger Jahre von Wilson und anderen die „continuous interleaved sampling“ (CIS) Strategie entwickelt ([Wilson et al., 1991](#) und [Wilson et al., 1993](#)). Bei dieser Strategie werden zu jeder Elektrode kurze Pulse in nichtüberlappenden Sequenzen gesandt. Das Signal wird durch eine Filterbank von Bandpassfiltern in bis zu zwölf Kanäle aufgeteilt. Die zwölf Kanäle definieren im weiteren Verlauf die Stimulationsstärke, die an den intrakochleären Elektroden in Form der biphasischen Pulse abgegeben wird. Durch zeitlichen Versatz dieser kurzen biphasischen Pulse können Überlappungen und Interaktionen zwischen den Kanälen minimiert werden (vgl. [Wilson et al., 1991](#)).

Da die weit verbreitete CIS-Strategie Restriktionen speziell bei der Unterscheidung von Tonhöhen und demzufolge bei der Wahrnehmung von Stimmqualitäten und prosodischen Mustern aufweist, wurde die „channel specific sampling sequences“ (CSSS) Strategie entwickelt. Bei dieser Methode wird die Feinstruktur mit Hilfe von kanalspezifischen Sequenzen übertragen (vgl. [Zierhofer, 2007](#) und [Schatzer et al., 2006](#)). Nach [Mitterbacher \(2004\)](#) wird bei der CIS-Strategie durch die Einhüllende der Filterausgangssignale die Stimulationsamplitude berechnet, wobei die in den Signalen enthaltene Phaseninformation – auch als Feinstruktur bezeichnet – verloren geht. Im Gegensatz dazu werden bei der neu entwickelten CSSS-Strategie bei jedem Nulldurchgang der Filterausgänge kanalspezifische Sequenzen gestartet. Dadurch wird die Feinstruktur im Stimulationssignal weitgehend erhalten.

Die neue CSSS-Strategie ist neben der CIS-Strategie im Kochleaimplantat Pulsar von Med-El integriert und verspricht eine verbesserte Perzeption prosodischer

Merkmale, da die Grundfrequenzverläufe besonders deutlich hörbar sein sollen. Bis zum gegenwärtigen Zeitpunkt gibt es noch keine aussagekräftigen Studien über diese Methode. Zur Überprüfung der Wahrnehmung der Prosodie existieren keine adäquaten Tests, wodurch S. Brill, der Leiter dieser Arbeit, die Idee zu dem in der Vorgängerarbeit (vgl. [Bauer, 2005](#)) entworfenen und in der vorliegenden Arbeit weiterentwickelten Prosodieverständnistest hatte.

### 1.2 Aktueller Wissensstand

Über die Prosodiewahrnehmung Normalhörender wird in der Literatur ausreichend berichtet. Studien über den Zusammenhang von syntaktisch mehrdeutigen Sätzen und der Prosodie liegen beispielsweise von [Lehiste \(1973\)](#) und [Kraljic und Brennan \(2005\)](#) vor, während u. a. [Darwin et al. \(2003\)](#), [Lehiste und Meltzer \(1973\)](#) und [Fu et al. \(2004\)](#) sich eingehend mit der Sprechererkennung beschäftigten. Andere Autoren führten Experimenten nach verschiedenen – nicht in direktem Bezug zur vorliegenden Arbeit stehenden – Fragestellungen durch. Untersucht wurde z. B. die Prosodiewahrnehmung in Relation zum Alter ([Kiss und Ennis, 2001](#)) oder im Zusammenhang mit musikalischer Begabung ([Magne et al., 2006](#)), Einflüsse der Prosodie beim Spracherwerb bei Kindern ([Shatzman und McQueen, 2006](#) und [Peppé und McCann, 2003](#)) oder die natürlich vorkommenden Variationen der Prosodie ([Peppé et al., 2000](#)). Es liegt somit eine Vielzahl von Studien über die Prosodie und deren Perzeption durch Normalhörende vor.

Vergleichsweise wenig Untersuchungen existieren hingegen mit hörgeschädigten Personen oder Hörhilfeträgern. [Leder et al. \(1986\)](#) testete die akustischen Parameter Grundfrequenz, Dauer und Intensität an CI-Trägern vor und nach der Implantation sowie im weiteren Verlauf. Als Vergleichsgruppe bediente er sich zum einen stark Ertaubter entsprechend der Hörsituation der Testgruppe vor der Implantation und zum anderen einer Gruppe Normalhörender. Bei den Cochleaimplantaten handelte es sich um Einkanalssysteme. [Waltzman und Hochberg \(1990\)](#) untersuchten Träger von Mehrkanalimplantaten mit Hilfe der SPAC-Batterie („Speech Pattern Contrasts“). Auch Boothroyd bediente sich in seinen Experimenten ([Boothroyd, 1984](#) und [Boothroyd, 1987](#)) der von ihm entwickelten SPAC-Batterie. In der letztgenannten Untersuchung testete er CI-Träger mit einkanaligen (House 3M) und mehrkanaligen (Nucleus 22) Implantaten. Hierbei schnitten die House 3M-Träger mit Erkennungsraten von durchschnittlich 90 % bei den Satzmodi und 96 % beim Sprechergeschlecht

etwas besser ab als die Nucleus 22-Patienten. [Richardson et al. \(1998\)](#), [Meister et al. \(2007b\)](#) und [Nie et al. \(2006\)](#) sind ebenfalls Autoren, die sich mit der Perzeption prosodischer Aspekte bei CI-Trägern beschäftigten.

Im englischsprachigen Raum dient die bereits erwähnte SPAC-Batterie als Instrumentarium zur Untersuchung der Prosodieperzeption. Diese berücksichtigt vier prosodische Aspekte: die Lokalisation von Satzakkzenten, die Bestimmung des Sprechergeschlechts, die Zuordnung von Satzmodi sowie die Unterscheidung natürlicher und monotoner Sprechweise. Letztgenanntes untersucht das Erkennen von Grundfrequenzänderungen der natürlich gesprochenen Sprache. Ferner werden außer den Suprasegmentaltests acht Subtests hinsichtlich der Unterscheidung segmentaler Merkmale durchgeführt (vgl. [Boothroyd, 1987](#)).

Des Weiteren steht die MAC-Batterie („Minimal Auditory Capabilities“), die seit den neunziger Jahren auch die deutsche Sprache abdeckt, mit einigen Ansätzen zur Evaluierung prosodischer Merkmale zur Verfügung (vgl. [Owens et al., 1982](#)). Die MAC-Batterie besteht aus 13 Hörtests und einem Lippenlesetest. Untersucht werden Satzmodus, Satzakkente und die Unterscheidung zwischen Stimme und Umweltgeräusch. Letztgenanntes erscheint relevant im Hinblick auf Nutzen und Komfort einer Hörhilfe. Würde statt einer Stimme stets ein Geräusch wahrgenommen, wäre die Hörhilfe unbefriedigend. Die übrigen Hörtests evaluieren die Erkennung von Phonemen oder Wörtern und haben keinen direkten Bezug zur Prosodie. Die deutschsprachige Version der MAC-Batterie umfasst die Testarten Satzmodus, Sprechergeschlecht sowie einen Satzidentifikations-Test zur Unterscheidung der Silbenanzahl in einem Satz.

2007 wurde für den deutschsprachigen Raum von [Meister et al. \(2007a\)](#) eine Prosodietestbatterie entwickelt, mit der die prosodischen Parameter Dauer, Satzmodus, Satzakkent und Geschlecht getestet werden können. Der Test gliedert sich in zwei Teile. Im ersten Teil sollen durch Darbietung natürlicher Stimuli von verschiedenen Sprechern alltägliche Hörsituationen simuliert werden. Im zweiten Teil werden durch Modifizierung der natürlichsprachigen Aufnahmen kleine Perzeptionsunterschiede hinsichtlich prosodischer Merkmale erfasst. Verändert wurden die akustischen Parameter Grundfrequenz und Vokaldauer. Bislang liegen allerdings ausschließlich Referenzmessungen mit einer kleineren Gruppe normalhörender Probanden vor, so dass Ergebnisse mit CI-Trägern abzuwarten bleiben.

Somit stehen für den deutschen Sprachgebrauch adäquate Testverfahren, die das Prosodieverständnis in ausreichendem Rahmen prüfen, derzeit nur in sehr begrenztem Maß zur Verfügung.

### 1.3 Fragestellung und Zielsetzung

Da technische Hörhilfen und insbesondere Kochleaimplantate einer stetigen Weiterentwicklung unterworfen sind und das Prosodieverständnis im alltäglichen Leben der Betroffenen eine herausragende Rolle spielt, ist die Entwicklung eines geeigneten Prosodietests von großer Relevanz.

Um die Eignung eines solchen Tests festzustellen, muss er vor Anwendung an Hörgeschädigten mit normalhörenden Personen evaluiert werden. Diese sollen später als Referenzgruppe dienen und Vergleichswerte liefern. Da in der Literatur schon zahlreiche Studien über das Prosodieverständnis Normalhörender in ruhiger Umgebung existieren, wurden in der vorliegenden Arbeit durch Überlagerung des Sprachmaterials mit einem unterschiedlich lauten Störgeräusch verschiedene Schwierigkeitsstufen erzeugt. So sollten unterschiedliche Hörkonditionen geschaffen werden, die das gesamte Leistungsspektrum von absolutem Raten bis hin zu fehlerfreiem Urteilen der Probanden abdeckten. Dies erscheint auch sinnvoll im Hinblick auf spätere Tests mit CI-Trägern. Nach [Dorman et al. \(1998b\)](#) und [Fu et al. \(1998\)](#) weisen die meisten kochleaimplantierten Patienten eine hohe Sprachwahrnehmung in ruhiger Umgebung auf, wohingegen sie in Gegenwart von Störgeräuschen deutlich schlechter abschneiden.

Mit der Wahrnehmung von Sprache unter dem Einfluss eines Störgeräuschs beschäftigten sich u. a. die Autoren [Stelmachowicz et al. \(1990\)](#), [Kalikow et al. \(1977\)](#), [Fu und Shannon \(1999\)](#) und [Bilger et al. \(1979\)](#). [Kalikow et al.](#) entwickelten einen Sprachverständnistest, bei dem die Testitems in „babble-speech“ präsentiert wurden. Der Studie zufolge ist die Überlagerung der Items mit dieser Art von Störlärm für die Versuchsteilnehmer schwieriger als nichtsprachliche Geräusche, da falsche Sprachhinweise gegeben werden. Prosodische Aspekte wurden allerdings nur insoweit berücksichtigt, als dass bei der Auswahl des Sprachmaterials auf eine Gleichverteilung phonetischer Faktoren geachtet wurde, um möglichst natürliche Sprache darzustellen. [Fu et al. \(1998\)](#) untersuchten u. a. an Normalhörenden die Fähigkeit, Vokale und Konsonanten unter variierenden Bedingungen korrekt zuzuordnen. Dabei wurden Schwierigkeitsgrade erzeugt, die Kochleaimplantate simulierten. Andere Studien, in denen ebenfalls die Simulation von Kochleaimplantaten zum Einsatz kamen, liegen von [Fu et al. \(2004\)](#), [Dorman et al. \(1998a\)](#) und [Dorman et al. \(1998b\)](#) vor. [Stelmachowicz et al. \(1990\)](#) verwendeten bei ihren Sprachtests sowohl ein Breitbandrauschen als auch gefiltertes Tiefpassrauschen als Hintergrundgeräusch. Der Grund für die Anwendung des Tiefpassrauschens ist den Autoren der Studie zufol-

ge in dem Umstand zu sehen, dass die meisten Umweltgeräusche hauptsächlich niederfrequente Energie enthalten, was bei Tests berücksichtigt werden sollte. Hörgeschädigte schnitten bei Darbietung des niederfrequenten Störlärms schlechter ab als bei einem Breitbandrauschen.

Einen anderen Ansatz zur Schaffung unterschiedlicher Schwierigkeitsgrade lieferten [Meister et al. \(2007a\)](#). In einer Prosodietestbatterie modifizierten sie natürlich-sprachliche Aufnahmen, indem sie die akustischen Parameter Grundfrequenz und Dauer mittels einer Sprachsignalverarbeitungssoftware veränderten. Auf diese Weise entstanden neun Stufen, die den Bereich zwischen den beiden Extrema (z. B. Aussage und Frage) abdeckten.

Die vorliegende Dissertation baut auf der Arbeit von [Bauer \(2005\)](#) auf, die ebenfalls unter der Leitung von S. Brill durchgeführt wurde. Im folgenden soll diese Vorgängerarbeit kurz erläutert werden:

Es existieren vier Untertests, in denen zwischen den jeweiligen Satzmodi, Prominenzen, Foki und Stimmqualitäten unterschieden werden soll. Für den Untertest Satzmodi wurden syntaktisch identische Sätze gewählt, die je nach Betonung der Aussage, der Frage oder dem Ausruf zuzuordnen sind. Beim zweiten Untertest Prominenz werden zwei syntaktisch gleiche Sätze in zwei Varianten dargeboten, die durch Betonung den Sinn des Gesagten unterscheiden. So kann beispielsweise der Satz „Sarah kauft Eimer und Schaufel und Handschuhe besorge ich“ zum einen bedeuten, dass Sarah Eimer und Handschuhe besorgt und zum anderen, dass die beiden letztgenannten Objekte von mir gekauft werden. Der Untertest Fokus ist relativ ähnlich: In den beiden syntaktisch gleichen Sätzen wird jeweils ein anderes Wort betont und dadurch der Sinn abgeändert. Bei der Unterscheidung der Stimmqualitäten werden die Sprecher Mann, Frau oder Kind dargeboten.

Die Vorgängerarbeit diente als „Proof-of-concept“ und wurde in entsprechend kleinem Rahmen an fünf Normalhörenden getestet. Basierend auf den genannten Untersuchungen wurde entschieden, sich in der vorliegenden Arbeit auf die beiden wichtigsten Subtests Satzmodus und Sprechergeschlecht zu beschränken und diese in größerem Rahmen und mit einer höheren Anzahl an Probanden durchzuführen.

Bei den Satzmodi ist der Grundfrequenzverlauf auf den letzten Silben eines Wortes das entscheidende Kriterium. Bei einer Aussage sinkt die Grundfrequenz gegen Ende des Satzes, während sie bei einer Frage ansteigt. Ein weiteres wichtiges Entscheidungsmerkmal liefert die Betonung als Ausdruck von Emotionen. Dies spielt speziell beim Ausruf, der stets als emotionsgeladen anzusehen ist, aber auch bei der Frage eine große Rolle. Beispielsweise kann ein Ausruf sowie eine Frage Staunen,



## 1 Einleitung

Ärger, Freude etc. ausdrücken. Die Betonung ist demnach ein wichtiges Kriterium zur Abgrenzung der Frage und des Ausrufs von der Aussage, die eher trocken und informationsvermittelnd ausgesprochen wird. Bei den der Betonung zugrunde liegenden akustischen Parametern handelt es sich um Amplituden- und Dauerveränderungen. Der Grundfrequenzverlauf ist bei betonten Silben ebenfalls wichtig, da die Stimme hier typischerweise angehoben wird.

Durch die Darbietung der Satzkategorien Aussage, Frage und Ausruf wurden folglich die drei Suprasegmentalia des Gegenstandsbereiches Prosodie, nämlich Intonation, Lautheit und Quantität (vgl. Möbius, 1993) abgedeckt, wobei der Schwerpunkt auf der Intonation lag. In den meisten Studien wurde hingegen auf die Satzkategorie Ausruf verzichtet und nur die hinsichtlich des Grundfrequenzverlaufs Extrema Aussage und Frage getestet.

Die Testart Sprecher geschlecht wurde ebenfalls beibehalten, da sie wichtige Aussagen über die Erkennung von Grundfrequenzen lieferte. Entgegen der Mehrzahl der Untersuchungen über die Unterscheidung von Stimmqualitäten, die sich auf die Sprecher Mann und Frau beschränkten, war in der hier vorliegenden sowie der Vorgängerarbeit als dritter Sprecher zusätzlich das Kind vertreten.

Folgende Fragestellungen sollten anhand des Tests untersucht werden:

- Wie gut nehmen Normalhörende prosodische Merkmale in unterschiedlich schwierigen Hörsituationen wahr? In welchem Bereich liegt ihr Leistungsspektrum?
- Sind die Leistungen der Probanden ähnlich oder treten größere Streuungen auf?
- Ist eine Testart schwerer als die andere?
- Treten Verwechslungen der Sprecher bzw. Kategorien auf und wenn ja welche?
- Gibt es Einflüsse des Satzinhaltes auf die Wahl der Satzkategorie und/oder der Sprecher? Welcher Art sind diese möglicherweise auftretenden Interaktionen?
- Ist das Satzmaterial geeignet oder treten Ausreißersätze auf?
- Stellt der Test per se ein adäquates Instrumentarium zur Beurteilung der Prosodieperzeption dar?



## 2 Versuchsplanung und Methodik

### 2.1 Auswahl der Probanden

Am Hauptversuch nahmen 30 normalhörende Erwachsene in einer Altersklasse zwischen 20 und 30 Jahren auf freiwilliger Basis teil. Als Kriterium für normales Hören wurde eine Hörschwelle von höchstens 20 dB im Frequenzbereich zwischen 250 Hz und 4 kHz und höchstens 30 dB bei 125 Hz und 8 kHz festgelegt.

Nach [Schumacher und Schulgen \(2007\)](#) müssen die Versuchsteilnehmer repräsentativ für die entsprechende Grundgesamtheit sein, um die Übertragbarkeit der Studienergebnisse auf andere Probanden (externe Validität) zu sichern. Mit Hilfe von Ein- und Ausschlusskriterien der Untersuchung wird die Zielpopulation festgelegt, auf die die Studienergebnisse verallgemeinert werden sollen. Da das Hauptkriterium Normalhörigkeit darstellte und sich im Laufe des Alters die Hörschwelle und damit eventuell auch das prosodische Sprachverständnis verändert, wurde die Altersklasse bewusst niedrig gewählt. Mit einer möglichen Änderung des Prosodieverständnisses in Relation zum Alter beschäftigten sich u. a. die Autoren [Kiss und Ennis \(2001\)](#) und [Wingfield et al. \(2000\)](#) und kamen zu teilweise unterschiedlichen Ergebnissen. Nach [Kiss und Ennis \(2001\)](#) sind alte Patienten signifikant schlechter in der Perzeption prosodischer Merkmale als jüngere gleichen Intellekts, wobei die altersbezogenen Unterschiede in der Wahrnehmung denen bei Memory-Tests ähneln. [Wingfield et al. \(2000\)](#) zufolge ist die Fähigkeit zur Prosodiewahrnehmung bei jungen und alten Testpersonen ähnlich, sofern Unterschiede bezüglich des Hörvermögens bedacht werden. Um dennoch eventuelle altersbedingte Einflüsse auszuschließen, wurde die genannte Altersklasse als Auswahlkriterium vorgegeben. Zur Überprüfung der Hörschwelle und somit eines normalen Gehörs wurde von jedem Probanden vor der eigentlichen Versuchsdurchführung ein Tonaudiogramm angefertigt.

Nach [Hellbrück et al. \(1984\)](#) kommen geschlechtsspezifische Unterschiede in der Lautstärkenempfindung nur bei einer Darbietung der Stimuli über Kopfhörer, die den Gehörgang von außen abdichten, zum Tragen. Dabei nimmt die Lautheit mit zunehmendem Gehörgangsvolumen ab, so dass Frauen, die im Mittel einen kleineren Gehörgang aufweisen als Männer, unter sonst gleichen Bedingungen lauter hören. Da

bei diesen Untersuchungen eine Darbietung der Stimuli im Freifeld stattfand, musste auf eine Gleichverteilung der Geschlechter nicht geachtet werden.

Am Hauptversuch nahmen 14 männliche und 16 weibliche Probanden teil.

### 2.2 Tonaudiogramm

Das Tonaudiogramm wurde mit dem Audiometer Westra CAD03/1 in Verbindung mit einem Computer und Westra-CAD03-Software in der Hörkabine B der HNO-Univ.-Klinik Würzburg erstellt, in der unter gleichen Bedingungen Patienten im Routinebetrieb untersucht werden. Der Gerätebeschreibung ([Westra GmbH](#)) zufolge handelt es sich um ein Reinton- und Sprachaudiometer nach IEC 645-1, das den Ton als gepulstes Signal darbietet. Es kann in einem Frequenzbereich von 125 Hz bis 12,5 kHz mit einer Frequenzgenauigkeit von 1 % gemessen werden. Der Lautstärkepegel erstreckt sich von –20 dB bis 120 dB. Ferner weist das Audiometer laut Hersteller einen hohen Bedienungskomfort auf.

Zur Aufzeichnung des Tonaudiogramms sitzt der Proband in einem separaten schalldichten Raum, der durch eine Glasscheibe Sichtkontakt zum Versuchsleiter ermöglicht. Das Signal wird durch Luftleitungskopfhörer (DT 48, Beyerdynamic) zunächst in allen Frequenzen für das rechte Ohr und anschließend für das linke Ohr dargeboten. Das jeweils nicht zu testende Ohr wird mit einem Schmalbandrauschen vertäubt. Angefangen bei –20 dB wird der Lautstärkepegel jeweils um 5 dB erhöht, bis der Proband das Signal hört. Daraufhin betätigt er eine Taste, wodurch der aktuelle Pegel abgespeichert wird. Für jede Frequenz wird der aufgezeichnete dB-Wert durch Wiederholung der Messung noch einmal bestätigt, bevor zur nächsten Frequenz übergegangen wird.

Vor jeder Messung müssen die Probandendaten in das von Westra-Electronic entwickelte zugehörige Patientenverwaltungsprogramm eingegeben werden. Es werden Name, Vorname, Geburtsdatum sowie Geschlecht in die vorgegebenen Felder eingetragen und gespeichert. Ohne diese Eingaben lässt sich die Messung nicht starten.

## 2.3 Aufbau der Testsätze

### 2.3.1 Verwendete Sprachdateien

Für den Test wurden die Originalsätze des im Rahmen einer Magisterarbeit (Bauer, 2005) erstellten Prosodietests verwendet. Die weibliche Stimme wurde von der Verfasserin der Magisterarbeit gesprochen. Die männliche Stimme sprach der Betreuer letztgenannter sowie der vorliegenden Arbeit, S. Brill. Beim kindlichen Sprecher handelte es sich um ein siebenjähriges, dialektfrei sprechendes Mädchen. Die Sätze waren mit 16 Bit und einer Abtastfrequenz von 44,1 kHz aufgezeichnet und im WAV-Dateiformat abgespeichert worden.

### 2.3.2 Störgeräusch

Um möglichst das gesamte Leistungsspektrum der Probanden mit dem Test erfassen zu können, wurden unterschiedliche Schwierigkeitsstufen erzeugt. Hierzu wurde ein Störgeräusch in verschiedenen Lautstärken verwendet. Die Einführung des Störgeräuschs ermöglichte ebenfalls die Bewertung des Prosodieverständnisses der Probanden unter simulierten Alltagsbedingungen. Dies basiert auf der Tatsache, dass Sprache im normalen Alltagsleben oft unter bestimmten Störeinflüssen wahrgenommen wird. Zu den in der Umwelt vorkommenden Geräuschen kommen bei Hörhilfeträgern noch geräteinterne Störquellen dazu.

Es wurde ein CCITT-Rauschen (Comité Consultatif International Télégraphique et Téléphonique) verwendet (vgl. International Telecommunication Union, 1988). Hierbei handelt es sich um ein Breitbandrauschen, dessen Frequenzspektrum dem gesprochenen Sprache nachgebildet ist. Es werden weitere Sprecher simuliert, wodurch sich dieses Störgeräusch somit gut für die Simulation normaler Alltagsbedingungen eignet. Die Schallpegelverteilung ist weitgehend konstant, so dass die Sätze gleichmäßig durch das Rauschen überlagert wurden.

### 2.3.3 Erzeugung verrauschter Satzdateien

Das CCITT-Rauschen wurde in unterschiedlichen Lautstärken den Sätzen überlagert. Abbildung 2.1 stellt zur Veranschaulichung die Erzeugung eines solchen verrauschten Satzes schematisch dar. Es werden sieben Einzelgraphiken (a)–(g) gezeigt, welche aufeinander aufbauen und in der letzten Graphik (g) im verrauschten Satz resultieren. Während auf der x-Achse der zeitliche Verlauf in Sekunden darge-

stellt ist, zeigen die y-Achsen der Einzelgraphiken (a), (c), (d), (f) und (g) jeweils die Auslenkung des Signals zwischen den Werten  $-1$  und  $1$ . Da in (b) und (e) hingegen der Aufbau der Geräuschpegel gezeigt werden soll, erstrecken sich die Werte der y-Achsen von  $0$  bei absoluter Stille bis  $1$  bei maximaler Intensität. Die sieben Einzelgraphiken sind durch vertikale Linien in die Zeitabschnitte A–G unterteilt.

Die drei ersten Graphiken veranschaulichen die Entstehung des Satzes, während in den drei folgenden Abbildungen die Erzeugung des Rauschens verdeutlicht wird. Die letzte Graphik zeigt eine Addition des Satzes (c) und des Rauschens (f), wodurch der verrauschte Satz (g) entsteht.

Abbildung (a) zeigt den ursprünglichen Satz, dessen Länge der Buchstabe D markiert. Würde man die Satzdatei so abspielen, sind Knacklaute zu Beginn und am Ende des Satzes möglich, die den Probanden irritieren und beeinflussen könnten. Aus diesem Grund wird der Satz über eine cosinusförmige Einblendrampe der Dauer von  $50$  ms eingeleitet und entsprechend über eine Ausblendrampe von ebenfalls  $50$  ms wieder ausgeleitet (b). In (c) ist der Satz inklusive der Ein- und Ausblendrampen C und E dargestellt. Die folgende Einzelgraphik (d) zeigt das CCITT-Rauschen im zeitlichen Verlauf. Mögliche störende Hintergrundgeräusche vor und nach der verrauschten Satzdatei sollten ausgeschaltet werden. Um dies zu erreichen, wurden jeweils  $50$  ms absolute Stille zu Beginn und am Ende des Rauschens eingefügt (A und G). Bei Betrachtung der Zeitachse fällt auf, dass das Rauschen länger ist als der Satz und diesen vollständig umfasst. Auf diese Weise sollten mögliche Hörspitzen des Satzes zu Beginn und am Ende überlagert werden sowie eine Gewöhnung der Probanden an das Rauschen mit dem Hinweis „Gleich geht es los“ erfolgen. Die Länge des Rauschens vor Satzbeginn und nach Satzende wurde auf jeweils  $300$  ms festgelegt. Im nächsten Schritt sollte das Rauschen ebenso wie der Satz in (b) durch Rampen ein- und ausgeblendet werden, um Knacklaute zu vermeiden. Aus Abbildung (e) geht hervor, dass Ein- sowie Ausblendrampe (B und F) mit einer Dauer von jeweils  $300$  ms ebenso lang sind wie das Rauschen vor und nach dem Satz. Der Vorteil dieser Methode ist ein bereits vollständiger Aufbau des Rauschens bei Satzbeginn und ein noch maximales Rauschen bei Satzende. Abbildung (f) stellt das erzeugte Rauschen (d) inklusive der Rampen (e) dar. In (g) ist die verrauschte Satzdatei als Summe des Satzes (c) und des Rauschens (f) abgebildet. Bei dem verrauschten Beispielsatz ( $+6$  dB SNR) handelt es sich um Satz Nr. 19 („Das Buch ist spannend“), der vom männlichen Sprecher als Frage gesprochen wurde.

Zur Visualisierung sind in den folgenden Abbildungen [2.2–2.7](#) der Zeitverlauf (jeweils obere Zeile) und ein Spektrogramm (jeweils untere Zeile) eines Satzes für die

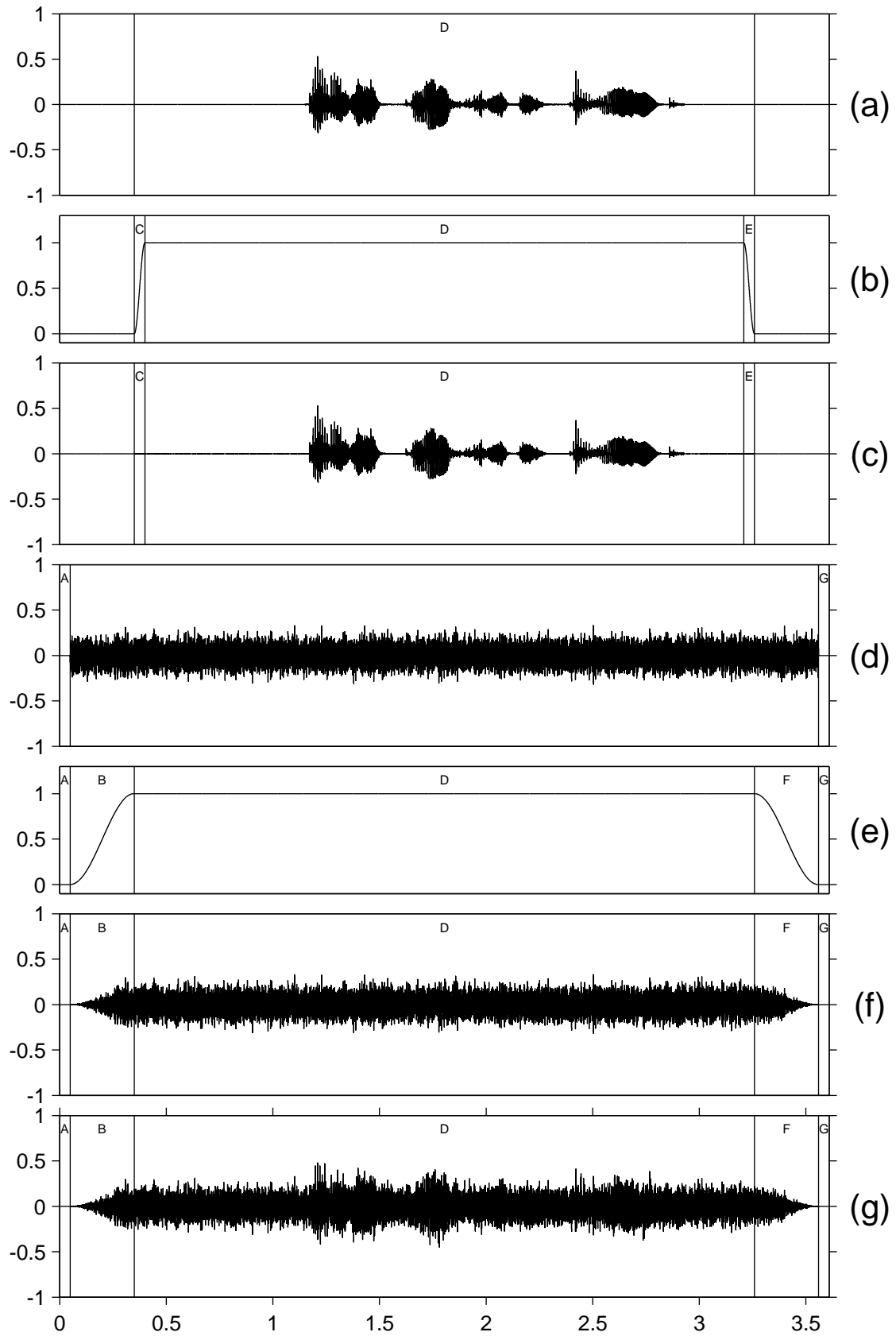


Abbildung 2.1: Erzeugung eines verrauschten Satzes

Sprecher Mann, Frau und Kind sowie die Kategorien Aussage, Frage und Ausruf dargestellt. Bei allen Graphiken wurde der Satz Nr. 1 („Da ist das Rathaus“) verwendet. Die x-Achse zeigt den am Computer geschnittenen relevanten Ausschnitt des Satzes in Sekunden. Auf der y-Achse der oberen Zeile (Zeitverlauf) ist die Signalauslenkung und auf der y-Achse des Spektrogramms die Frequenz in Hertz aufgetragen.

Bei der Erkennung der unterschiedlichen Sprecher und Kategorien ist die Tonhöhe bzw. Grundfrequenz ein wesentliches Kriterium. Während bei den Sprechern die absolute Tonhöhe, die Grundfrequenz per se, entscheidend ist, ist bei der Testart Kategorie der Verlauf der Grundfrequenz Hauptentscheidungsmerkmal. In den Abbildungen 2.2–2.7 ist die Grundfrequenz in den Spektrogrammen als helle Linie dargestellt. Die Grundfrequenz ist typischerweise nur bei stimmhaften Anteilen in der Sprache feststellbar und kann bei stimmlosen Lauten nicht ermittelt werden. Aus diesem Grund sind Zischlaute im Grundfrequenzverlauf der Spektrogramme nicht dargestellt.

Bei Betrachtung der Abbildungen 2.2, 2.3 und 2.4 werden die unterschiedlichen Grundfrequenzen der Sprechstimmen sehr deutlich. Der Mann spricht mit einer durchschnittlichen Frequenz von 96,45 Hz und ist im unteren Bereich des Spektrogramms positioniert. Die Stimme der Frau liegt mit durchschnittlich 190,7 Hz wesentlich oberhalb der des Mannes. Die kindliche Stimme besitzt mit einer durchschnittlichen Grundfrequenz von 225,3 Hz die höchste Stimmlage.

In Abbildung 2.5 wird der für die Aussage typische gegen Ende des Satzes abfallende Grundfrequenzverlauf erkennbar. Bei der Frage hingegen ist die Grundfrequenz durch einen ansteigenden Verlauf gekennzeichnet. Der Ruf kann durch Betonungen und Satzakkente zugeordnet werden, die meist mit erhöhter Grundfrequenz gesprochen werden. So entspricht die erhöhte Grundfrequenz zu Satzbeginn in Abbildung 2.7 dem betonten Wort „da“ des zugehörigen Satzes „Da ist das Rathaus!“.

### 2.4 Vorversuche

In der vorliegenden Arbeit sollte untersucht werden, wie gut das prosodische Sprachverständnis bei Normalhörenden in unterschiedlich schwierigen Hörsituationen ist. Zur Schaffung verschiedener Hörkonditionen wurde das Rauschen in seiner Lautheit modifiziert, während der Sprachstimulus konstant gehalten wurde. Das variierende Verhältnis des Rauschens zum Sprachsignal wurde durch den SNR (signal to noi-

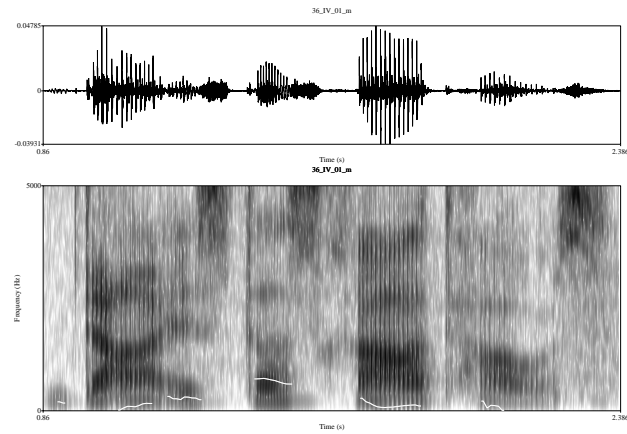


Abbildung 2.2: Zeitverlauf und Spektrogramm des männlichen Sprechers

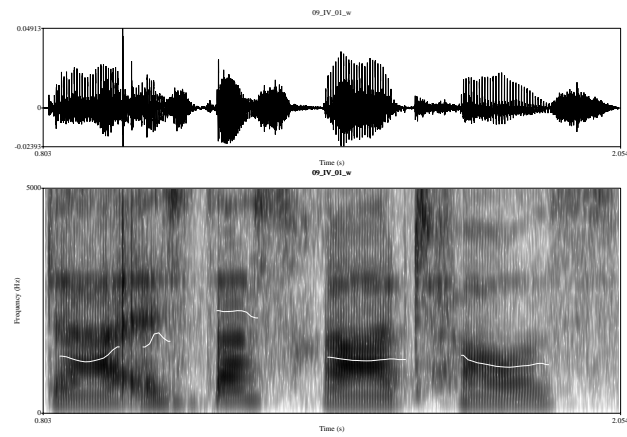


Abbildung 2.3: Zeitverlauf und Spektrogramm der weiblichen Sprecherin

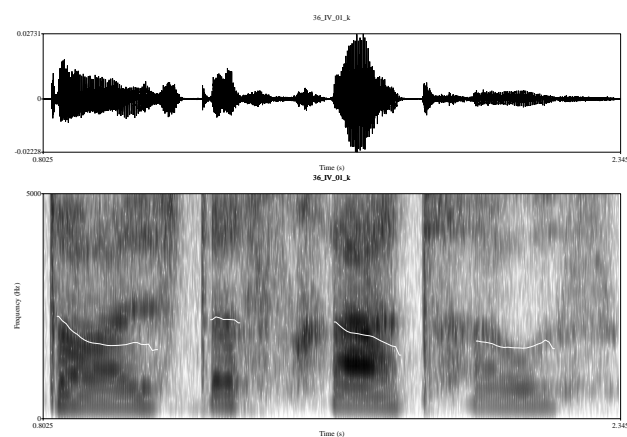


Abbildung 2.4: Zeitverlauf und Spektrogramm der kindlichen Sprecherin

## 2 Versuchsplanung und Methodik

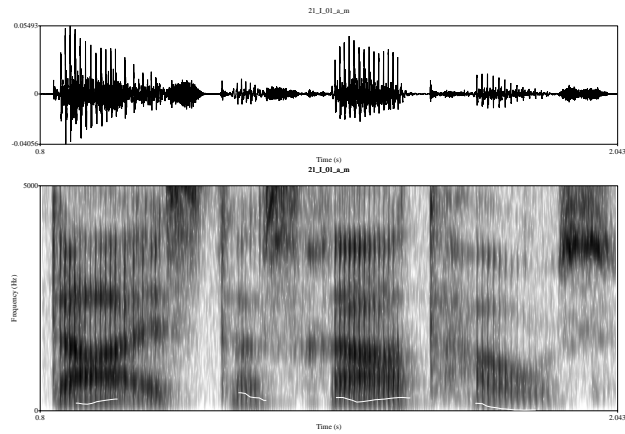


Abbildung 2.5: Zeitverlauf und Spektrogramm der Kategorie Aussage

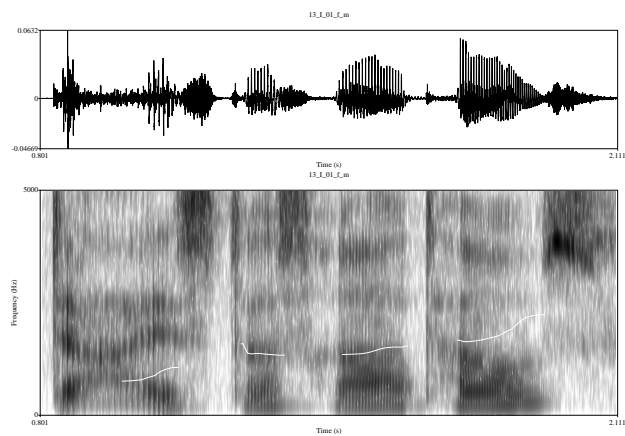


Abbildung 2.6: Zeitverlauf und Spektrogramm der Kategorie Frage

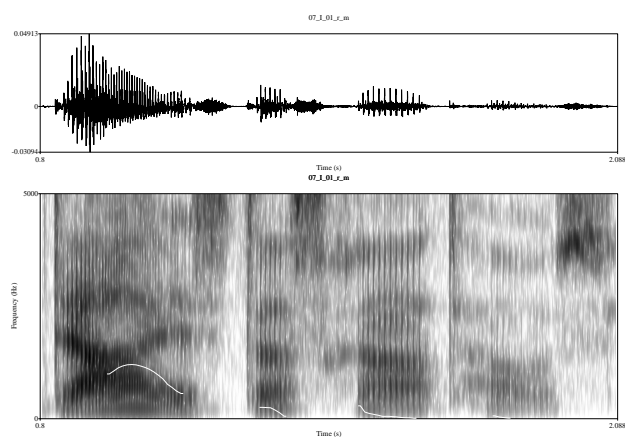


Abbildung 2.7: Zeitverlauf und Spektrogramm der Kategorie Ausruf



se ratio) festgelegt. Ein SNR von beispielsweise +5 dB bedeutete folglich ein relativ leises Rauschen, ein SNR von –15 dB ein relativ lautes Rauschen.

Da pro Testart jeweils drei Antwortmöglichkeiten zur Auswahl stehen, würde der Proband theoretisch allein durch reines Raten zu 33,3% korrekte Antworten geben. Bei nicht vorhandenem Rauschen hingegen müsste der Proband in der Lage sein, zu annähernd 100% richtige Urteile abzugeben. Demzufolge sollten durch Vorversuche diejenigen Hörkonditionen herausgefunden werden, die den Bereich zwischen reinem Raten und problemlosem Verstehen abdecken. Ein ausgewählter Bereich mit den Grenzwerten +10 dB SNR und –20 dB SNR wurde in Abständen von 1 dB auf eine DVD gebrannt und probegehört. Dies deckt sich mit Angaben aus der Literatur. [Darwin et al. \(2003\)](#) z. B. präsentierten ebenfalls Sätze unter sechs verschiedenen SNRs zwischen 0 und –15 dB. [Fu et al. \(1998\)](#) kombinierten in ihren Untersuchungen das Sprachsignal mit einem Störgeräusch für SNR-Werte von +24 dB bis –15 dB. Allerdings bleibt die Unterschiedlichkeit der Tests zu berücksichtigen. Fu testete die Perzeption von Vokalen und Konsonanten, während in der vorliegenden Arbeit vollständige Sätze verwendet werden. Nach dem Probehören erschien die leichteste Hörbedingung (+10 dB SNR) für Normalhörende und die vorliegende Arbeit als zu einfach. Die Hörkonditionen wurden deshalb auf einen Bereich zwischen +5 dB SNR und –15 dB SNR festgelegt. Im Rahmen der Vorversuche wurden die SNRs in einem Abstand von 5 dB präsentiert, so dass jeweils fünf Hörkonditionen pro Testart zur Verfügung standen. Die Hörkondition +10 dB SNR sollte am Anfang des jeweiligen Tests als leichter Übungsdurchlauf zur Verfügung stehen.

Es wurden sechs Probanden – vier männliche und zwei weibliche – für die Vorversuche ausgewählt.

### 2.4.1 Erstellung und Kombination der Listen

Benötigt wurden fünf Listen mit Testsätzen für die entsprechenden Hörkonditionen (+5 dB SNR, 0 dB SNR, –5 dB SNR, –10 dB SNR, –15 dB SNR) sowie eine Liste für einen dem Test vorangehenden leichten Übungsdurchlauf (+10 dB SNR). Jede Liste sollte 36 Sätze enthalten. Die Sätze der drei schon bestehenden Listen (vgl. [Bauer, 2005](#)) wurden hierzu mit dem Computer nach dem Zufallsprinzip neu verwürfelt. Durch die Randomisierung wurden drei neue Listen gewonnen, so dass insgesamt nun sechs unterschiedliche Listen zur Verfügung standen.

Nach [Wink \(2006\)](#), [Machin und Campbell \(2005\)](#) und [Schumacher und Schulgen \(2007\)](#) ist die Randomisierung ein wesentliches Kriterium zur Beurteilung der Qua-

lität einer Studie, da hierdurch die Voraussetzung für eine Vergleichbarkeit der Ergebnisse geschaffen wird. Dabei müssen per definitionem alle Einflussgrößen außer dem zu untersuchenden Einflussfaktor identisch sein. Die Forderung nach einer Fixierung aller anderen Einflussgrößen scheint jedoch unerfüllbar, da Patienten bzw. Probanden stets in so vielen bekannten und unbekanntem Faktoren variieren, dass es somit unmöglich ist, diese alle bei der Auswertung zu berücksichtigen (Schumacher und Schulgen, 2007). Um diesem Problem zu entgehen, wird seit den vierziger und fünfziger Jahren die Randomisierung (interne Validität) angewendet, die es ermöglicht, Ergebnisse auch tatsächlich dem zu untersuchenden Kriterium zuzuordnen.

Des Weiteren soll durch die Randomisierung ebenfalls vermieden werden, dass systematische Einflüsse („Bias“) die Ergebnisse verfälschen. Auch ist sie notwendig, um valide statistische Tests anwenden zu können (Wink, 2006).

Um die fünf Hörkonditionen den Listen entsprechend oben genannten Kriterien zuordnen zu können, erfolgte eine Randomisierung mit Hilfe des Münzwurfs. Hierdurch wurde gewährleistet, dass die Probanden nicht durch gleichmäßiges An- oder Absteigen der SNRs beeinflusst wurden und besser oder schlechter abschnitten als sie eigentlich waren.

Auch unter den Probanden wurde die Reihenfolge der Listen und die Kombination Liste — Hörkondition verändert, um Reihenfolgeeffekte zu minimieren. Hierzu wurden die Listen durchrotiert und jeweils beim nächsten Probanden mit der zweiten Liste des vorigen Probanden begonnen. Bei den Hörkonditionen wurde hingegen mit der letzten Hörkondition des vorigen Probanden begonnen und die Hörbedingungen somit in rückwärtiger Reihenfolge durchrotiert. Da durch dieses Verfahren der sechste Proband bei der Zuordnung der fünf Hörkonditionen zu den Listen wieder in Übereinstimmung mit dem ersten Probanden lag, wurden bei diesem die Hörkonditionen des ersten Probanden in rückwärtiger Reihenfolge verwendet. Beim Sprechertest wurde die zunächst für den Kategorientest festgelegte Reihenfolge der Hörkonditionen beibehalten. Es wurde jedoch mit der letzten Hörkondition des Kategorientests begonnen. Auf diese Weise wurden jedem Versuchsteilnehmer andere Kombinationen und Reihenfolgen präsentiert, wodurch unerwünschte Einflüsse ausgeschaltet werden sollten.

### 2.4.2 Festlegung der absoluten Lautheitspegel

Um eine Hörschädigung der Probanden nachweisbar sicher ausschließen zu können, musste der absolute Lautheitspegel in der Hörkabine C festgelegt werden. Dazu wurden am Audiometer verschiedene Dezibelwerte eingestellt und die ersten zwölf verrauschten Satzdateien der Liste A des Kategorientests in allen SNRs abgespielt. Beim Schallpegelmesser handelte es sich um einen 2260 Investigator der Firma Brüel & Kjær (SPL: A-bewertet, „fast“).

An der Probandenposition wurden dann die absoluten Lautheitspegel aufgezeichnet, wobei der lauteste Wert des jeweiligen Satzes notiert wurde. Durch eine Festlegung der Sprache auf 65 dB und additives Rauschen von maximal 15 dB waren Werte um 80 dB für die lauteste Hörkondition erstrebenswert. Da die Geräuschpegelgrenze für empfundene Lärmbelastigung bis hin zur Hörschädigung bei 90 dB liegt, hätte man so noch 10 dB Spielraum, falls sich nach den Vorversuchen herausstellen sollte, dass man doch auf ein lauterer Rauschen (z.B. -20 dB SNR oder sogar -25 dB SNR) zurückgreifen müsste. Das wäre dann der Fall, wenn die Probanden bei der in den Vorversuchen lautesten Hörkondition immer noch mehr richtige Ergebnisse hätten als dem Zufallsniveau der Antwortmöglichkeiten von 33,3% entspräche.

Es stellte sich heraus, dass bei Einstellung des Audiometers auf nominell 95 dB die absoluten Lautheitspegel für die lauteste Hörkondition bei 80 dB lagen. Der höchste Wert lag hierbei für -15 dB SNR bei 80,2 dB und fiel für die übrigen SNRs weiter unter diesen Wert bis hin zu 57,1 dB bei +10 dB SNR (vgl. Tabelle 2.1). Bei Betrachtung der Tabelle 2.1 fällt auf, dass für die Hörkonditionen mit dem leiseren Störgeräusch stärkere Signalschwankungen vorliegen als für die lauterer. Dies erklärt sich folgendermaßen: Im Bereich von -15 dB SNR bis -5 dB SNR dominiert das Rauschen. Daraus resultieren gleichmäßigere Werte und plausible Pegelabstände. Das Signal ist relativ konstant und leicht zu messen. Bei 0 dB SNR bis +10 dB SNR dominiert die Sprache. Die in Sprache vorhandenen Signalspitzen verursachen unterschiedliche Werte und damit auch größere Signalschwankungen.

### 2.4.3 Auswertung der Vorversuche

Für einen ersten Überblick wurde die Tabelle 2.2 erstellt, die die Antworten der sechs Probanden für den jeweiligen SNR zusammenfasst: Daraus ist ersichtlich, dass die Probanden sowohl im Kategorien- als auch im Sprechertest besser abschnitten als erwartet. Schon bei 0 dB SNR wurden 100% korrekte Urteile im Sprechertest bzw.

Tabelle 2.1: Schallpegel (Kabine C) bei nominell 95 dB (Audiometeranzeige) für die Liste A; gemessen in 1,0 m Abstand

| -15dB<br>SNR | -10dB<br>SNR | -5dB<br>SNR | 0dB SNR | 5dB SNR | 10dB<br>SNR |
|--------------|--------------|-------------|---------|---------|-------------|
| 79,9         | 75,1         | 70,3        | 65,7    | 63,0    | 60,8        |
| 80,2         | 75,3         | 70,6        | 66,4    | 63,5    | 62,1        |
| 80,2         | 75,3         | 70,8        | 66,7    | 64,7    | 62,8        |
| 80,0         | 75,2         | 70,2        | 66,0    | 62,5    | 60,8        |
| 80,2         | 75,7         | 70,1        | 66,3    | 63,2    | 62,1        |
| 79,9         | 75,2         | 70,6        | 66,6    | 62,5    | 60,4        |
| 79,9         | 75,1         | 70,1        | 65,3    | 61,2    | 68,3        |
| 80,2         | 75,3         | 70,5        | 66,1    | 62,6    | 60,4        |
| 80,1         | 75,3         | 70,6        | 66,5    | 64,1    | 62,8        |
| 80,1         | 75,3         | 70,5        | 66,1    | 62,3    | 59,5        |
| 80,1         | 75,3         | 70,6        | 66,2    | 63,0    | 61,9        |
| 80,1         | 75,0         | 70,1        | 65,3    | 60,9    | 57,1        |

Tabelle 2.2: Gesamtergebnisse der Vorgruppe: Anzahl der richtigen Antworten für den jeweiligen SNR

| SNR | Sprecher<br>(Treffer) | Sprecher<br>(%) | Kategorie<br>(Treffer) | Kategorie<br>(%) |
|-----|-----------------------|-----------------|------------------------|------------------|
| 5   | 35,8                  | 99,5            | 34,8                   | 96,8             |
| 0   | 36,0                  | 100,0           | 35,3                   | 98,2             |
| -5  | 34,8                  | 96,8            | 32,8                   | 91,2             |
| -10 | 30,0                  | 83,3            | 26,7                   | 74,1             |
| -15 | 22,2                  | 61,6            | 19,0                   | 52,8             |

98,15% im Kategorientest erreicht. Die Probanden waren sogar besser als bei +5 dB SNR, der leichtesten Hörkondition, was allerdings bei den wenigen falschen Antworten eher für Interpretationsfehler spricht. Die Werte der schwersten Hörkondition –15 dB SNR lagen beträchtlich über der 33,3%-Grenze der reinen Ratewahrscheinlichkeit. Es wurden 61,57% im Sprechertest und 52,78% im Kategorientest erzielt, so dass angenommen werden muss, dass die Probanden noch immer genügend Komponenten des Sprachsignals wahrnehmen konnten, um häufig die richtige Antwort zu geben.

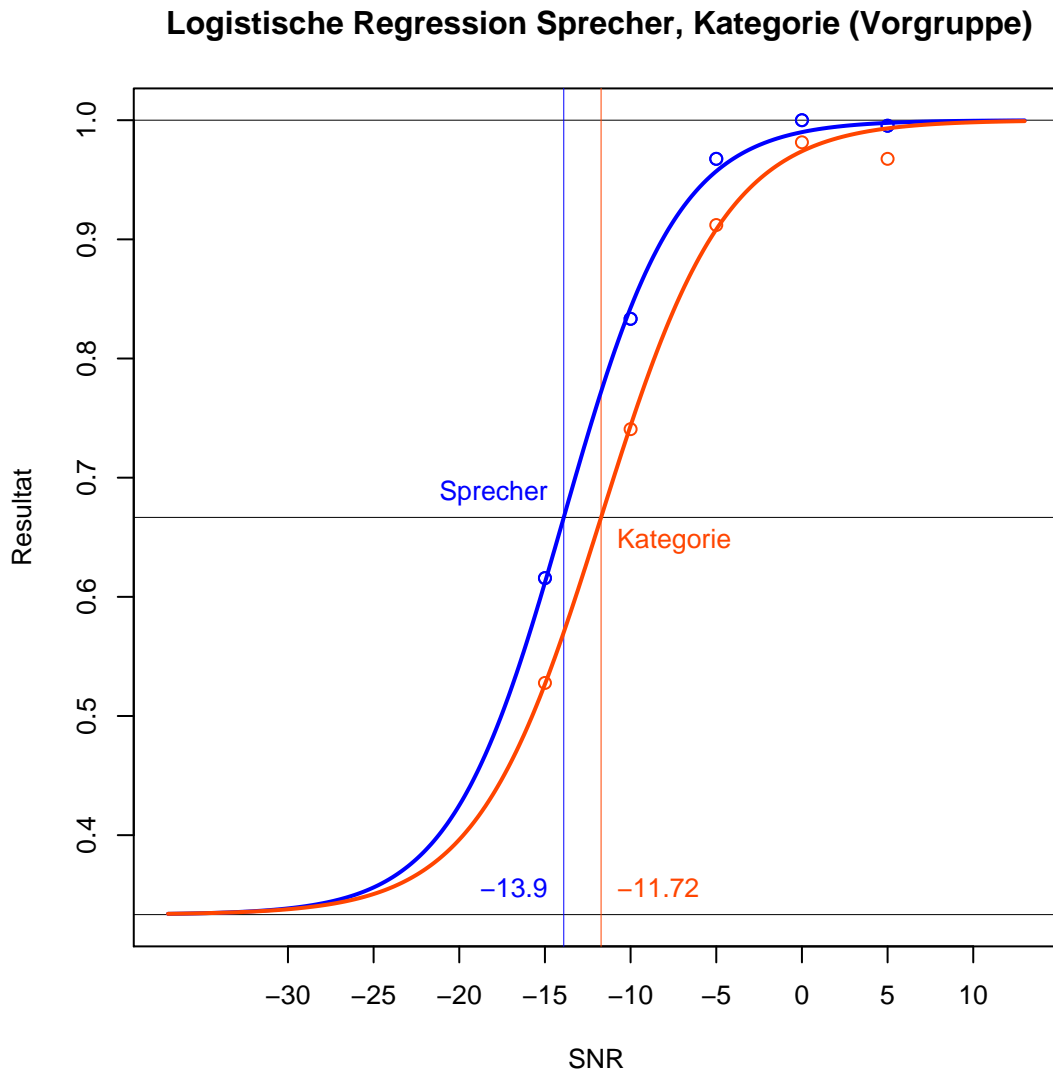


Abbildung 2.8: Resultate der Vorgruppe für die jeweiligen SNRs für Sprecher- und Kategorientest

Um entscheiden zu können, inwieweit die Tests aufgrund der Ergebnisse der Vorgruppe für die Hauptgruppe erschwert und modifiziert werden müssten, wurden logistische Regressionen (s. Abschnitt 2.9) berechnet, die in den Abbildungen 2.8, 2.9 und 2.10 dargestellt sind. Abbildung 2.8 fasst die Antworten der sechs Probanden in jeweils einer logistischen Regression für jede Testart zusammen. Die Graphik zeigt den Wendepunkt der Kurve für den Sprechertest bei  $-13,82$  dB SNR und für den Kategorientest bei  $-11,91$  dB SNR an. Demzufolge sind an den genannten Punkten immer noch die Hälfte aller Antworten korrekt und der Test muss folglich erschwert wer-

den. Außerdem ist bereits erkennbar, dass der Sprechertest den Probanden leichter zu fallen scheint als der Kategorientest.

Da die Konfidenzintervalle der logistischen Regressionen in Abbildung 2.8 jedoch zu groß sind, um genauere Aussagen bezüglich der Modifikation des Tests für die Hauptgruppe treffen zu können, wurden die individuellen Kurvenverläufe für alle Probanden in den Graphiken 2.9 und 2.10 dargestellt.

Die individuellen Kurvenverläufe sollen zuerst für den Sprechertest näher betrachtet werden (s. Abbildung 2.9). Auffallend ist, dass die Spanne der unterschiedlichen Steigungen sehr weitreichend ist. Bei manchen Probanden war ein flacherer Verlauf der Kurve zu verzeichnen als bei anderen, woraus ersichtlich ist, dass bei manchen Probanden der Übergang zwischen 33,3% und 100% der richtigen Antworten auf eine Spannweite von nur 10 dB begrenzt ist, während sich bei anderen die Spannweite auf bis zu 25 dB erstreckt. Der Graphik zufolge wäre also für den Sprechertest eine Ausdehnung von ca. 25 dB in einem Bereich von 0 dB SNR bis ca. -25 dB SNR erstrebenswert.

In Graphik 2.10 wird die Streuung der Antworten der verschiedenen Probanden für den Kategorientest deutlich. Während bei 0 dB SNR die Anzahl der richtigen Antworten bei allen Probanden noch annähernd 100% beträgt, liegt der Wendepunkt der Kurve (66,6%) zwischen -8 dB SNR des schlechtesten und -14 dB SNR des besten Probanden. Die Annäherung an die 33,3%-Grenze der reinen Ratewahrscheinlichkeit scheint zwischen -18 dB SNR und -25 dB SNR zu liegen.

### 2.5 Hauptversuche

Basierend auf den Ergebnissen der Vorversuche wurde für die Hauptversuche eine Spannweite der SNRs von 24 dB festgelegt. Die Abstände zwischen den jeweiligen SNRs sollten 6 dB betragen, so dass die Hörkonditionen -24 dB SNR, -18 dB SNR, -12 dB SNR, -6 dB SNR und 0 dB SNR resultierten. Der Übungsdurchlauf wurde auf +6 dB SNR festgelegt. Diese Kriterien sollten sowohl für den Sprecher- als auch für den Kategorientest gelten, was zudem die Versuchsdurchführung vereinfachen sollte.

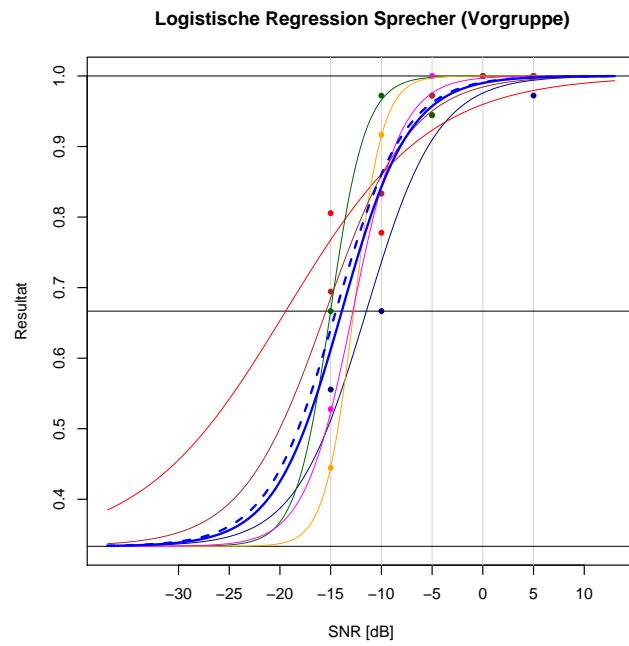


Abbildung 2.9: Resultate aller Probanden der Vorgruppe für den Sprechertest

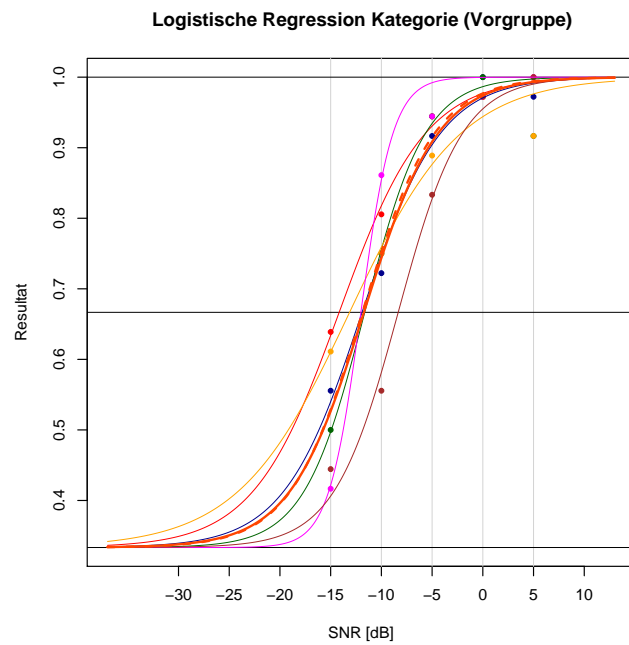


Abbildung 2.10: Resultate aller Probanden der Vorgruppe für den Kategorientest

### 2.5.1 Kombination der Listen

Da an den Hauptversuchen 30 Probanden teilnehmen sollten, war eine erneute Zuordnung der Hörkonditionen zu den Listen und den Probanden nötig. Die Anzahl der Probanden erklärt sich damit, dass es sich um fünf unterschiedliche Hörkonditionen, um sechs unterschiedliche Listen sowie um die zwei Testarten (Sprecher- und Kategorientest) handelt. Um sowohl Hörbedingungen, Listen als auch Testart gleichmäßig auf die Probanden zu verteilen und zu gewährleisten, dass diese in den Hauptversuchen genau gleich oft vorkamen, waren als kleinstes gemeinsames Vielfaches 30 Probanden nötig.

Es galt nun, zunächst eine Reihenfolge für die fünf unterschiedlich schweren Hörkonditionen festzulegen, um den zeitlichen Verlauf des Versuchs zu berücksichtigen. Dadurch sollten etwaige Anfangsschwierigkeiten zu Beginn des Versuchs, mögliche Lernprozesse oder eventuell mangelnde Motivation und Konzentration gegen Ende des Versuchs auf alle Hörkonditionen gleichmäßig verteilt werden. So sollte vermieden werden, dass eine Hörkondition z. B. häufiger am Anfang des Tests vorkam und somit den etwaigen Anfangsschwierigkeiten des Probanden häufiger ausgesetzt war als die anderen Hörkonditionen, wodurch sie schwieriger erscheinen würde als sie eigentlich war. Prinzipiell existierten 120 Möglichkeiten, die Reihenfolge der Hörkonditionen anzuordnen. Die Reihenfolge für die erste Testperson wurde mit dem Computer randomisiert. Tabelle 2.3 zeigt dies in der dritten Zeile (LMJIK für den Sprechertest). Für die folgenden Probanden wurde diese Reihenfolge durchrotiert, wobei jeweils mit der letzten Hörkondition des vorigen Probanden begonnen wurde. Die Reihenfolge der Hörkonditionen im Sprechertest war demnach für den zweiten Probanden KLMJI. Durch diese Rotation war nach fünf Probanden automatisch die erste Hörkondition des ersten Probanden wieder an erster Stelle, so dass jeweils nach fünf Probanden neu randomisiert werden musste. Danach wurde erneut durchrotiert. So erhielt jeder Proband eine andere Reihenfolge und die unterschiedliche Schwierigkeit der Hörkonditionen konnte durch den zeitlichen Verlauf des Tests nicht beeinflusst werden.

Bei der Verteilung der Listen auf die Hörkonditionen musste berücksichtigt werden, dass die Listen unterschiedlich schwer sein könnten. Um dies bei der Auswertung der Hauptversuche feststellen zu können, wurden die Listen mit jeweils wechselnden Hörkonditionen kombiniert. Kombinierte man beispielsweise stets Liste A mit der schwierigsten Hörkondition und Liste B mit der leichtesten Hörkondition, könnte man nicht auswerten, ob nun Liste A oder Liste B schwieriger ist oder kein Unter-



schied besteht. Die sechste Liste wurde jeweils mit der Hörkondition +6 dB SNR für den Übungsdurchlauf am Anfang des Tests verwendet.

Um eine Zwischenauswertung des Versuchs zu ermöglichen, wurden die Probanden in Fünferblocks aufgeteilt. Die ersten fünf Probanden begannen mit dem Sprechertest, während die nächsten fünf mit dem Kategorientest anfangen usw. Auf diese Weise wurde eine systematische Bevorzugung entweder des Sprecher- oder des Kategorientests ausgeschlossen. Etwaige Anfangsschwierigkeiten oder Lernprozesse der Probanden würden so keine der Testarten benachteiligen oder begünstigen. Nach jeweils zehn Probanden könnte eine Zwischenauswertung durchgeführt werden, da beide Testarten im zeitlichen Verlauf dann gleich verteilt sind. Die Einteilung in Fünferblocks ergibt auch im Hinblick auf die Reihenfolge der Hörkonditionen Sinn, da diese immer im Fünftakt durchrotiert wurden. Bei diesem Vorgehen handelte es sich also um eine „permutierte Blockrandomisierung“.

Zur Veranschaulichung zeigt Tabelle 2.3 einen Ausschnitt der Kombination Testart – Liste – Hörkondition für die ersten fünf Probanden. Die beiden Spalten der Tabelle stehen für den Sprecher (S)- bzw. den Kategorientest (K). Jeweils drei Zeilen sind einem Probanden zuzuordnen. Die jeweils erste Zeile bezeichnet Sprecher- bzw. Kategorientest, die zweite Zeile die Listen und die dritte Zeile die Hörkonditionen. Die Hörkonditionen wurden der Übersichtlichkeit halber durch die Buchstaben I (für –24 dB SNR) – M (für 0 dB SNR) ersetzt.

### 2.5.2 Festlegung der absoluten Lautheitspegel

Da das Rauschen für die Hauptversuche für die lauteste Hörkondition auf –24 dB SNR angehoben wurde, mussten die Schallpegel in der Kabine C erneut gemessen werden. Um eine Hörschädigung sicher auszuschließen, wurde der Sprachstimulus auf 60 dB abgesenkt. Dadurch sollte durch das um maximal 9 dB angehobene Rauschen der Schallpegel des Gesamtstimulus bei etwa 84 dB liegen. Es stellte sich heraus, dass bei einer Einstellung des Audiometers auf 91 dB für die Hörkondition –24 dB SNR Werte von höchstens 84,4 dB vorlagen. Die Lautheitspegel wurden beim Sprechertest für die Liste A in allen Hörkonditionen und beim Kategorientest für die Liste A in der lautesten Hörkondition aufgezeichnet (s. Tabellen 2.4 und 2.5).

## 2 Versuchsplanung und Methodik

Tabelle 2.3: Kombination der Testarten, Listen und Hörkonditionen für die ersten fünf Probanden

S, K : Sprecher-, Kategorientest

A–F : Listen

I–M : Hörkonditionen (–24 dB SNR – 0 dB SNR)

|     |   |   |   |   |   |   |   |   |   |   |
|-----|---|---|---|---|---|---|---|---|---|---|
| P01 | S | S | S | S | S | K | K | K | K | K |
|     | A | C | E | B | D | D | B | F | C | E |
|     | L | M | J | I | K | M | L | J | I | K |
| P02 | S | S | S | S | S | K | K | K | K | K |
|     | F | C | E | A | D | A | F | D | B | E |
|     | K | L | M | J | I | K | M | L | J | I |
| P03 | S | S | S | S | S | K | K | K | K | K |
|     | F | B | E | A | C | A | C | B | F | D |
|     | I | K | L | M | J | I | K | M | L | J |
| P04 | S | S | S | S | S | K | K | K | K | K |
|     | E | B | D | A | C | F | C | E | D | B |
|     | J | I | K | L | M | J | I | K | M | L |
| P05 | S | S | S | S | S | K | K | K | K | K |
|     | E | A | D | F | C | D | B | E | A | F |
|     | M | J | I | K | L | L | J | I | K | M |

Tabelle 2.4: gemessene Schallpegel (Kabine C) bei nominell 91 dB (Audiometeranzeige) für die Liste A Sprecher

| -24 dB<br>SNR | -18 dB<br>SNR | -12 dB<br>SNR | -6 dB<br>SNR | 0 dB<br>SNR | 6 dB<br>SNR |
|---------------|---------------|---------------|--------------|-------------|-------------|
| 84,2          | 78,2          | 72,4          | 66,9         | 62,4        | 59,6        |
| 84,2          | 78,1          | 72,3          | 67,0         | 62,9        | 61,1        |
| 84,1          | 78,1          | 72,3          | 66,9         | 62,6        | 60,5        |
| 84,2          | 78,2          | 72,2          | 67,1         | 63,4        | 61,7        |
| 84,1          | 78,3          | 72,1          | 66,4         | 61,7        | 59,1        |
| 84,2          | 78,2          | 72,3          | 66,8         | 62,5        | 60,0        |
| 84,4          | 78,3          | 72,3          | 66,8         | 62,9        | 61,3        |
| 84,1          | 78,3          | 72,3          | 66,7         | 63,0        | 61,2        |
| 84,2          | 78,1          | 72,3          | 66,4         | 61,5        | 58,6        |
| 84,2          | 78,2          | 72,4          | 66,9         | 62,5        | 60,2        |
| 84,2          | 78,1          | 72,3          | 66,8         | 62,4        | 60,7        |
| 84,2          | 78,1          | 72,4          | 66,4         | 64,5        | 63,2        |

Tabelle 2.5: gemessene Schallpegel (Kabine C) bei nominell 91 dB (Audiometeranzeige) für die Liste A Kategorie

-24 dB SNR

84,2

84,2

84,1

84,2

84,4

84,2

84,2

84,2

84,2

84,2

84,2

84,2

## 2.6 Versuchsdurchführung

### 2.6.1 Freifelddarbietung im schallgedämpften Raum

Sowohl Vor- als auch Hauptversuche wurden in Form der Freifelddarbietung in der Hörkabine C der HNO-Univ.-Klinik Würzburg durchgeführt. Bei der Freifelddarbietung muss darauf geachtet werden, dass der Raum sehr reflexionsarm ist, um mögliche Störungen durch Nebengeräusche oder andere Einflussfaktoren auszuschalten. Auch wären in einem nicht schallgedämpften Raum Reflexionen der Schallquellen hörbar, die die Probanden in ihren Antworten beeinflussen könnten. Die Hörkabine C war bei ihrem Bau durch schalldämpfende Materialien in Wänden, Boden und Decke sowie einer ebenfalls schallabsorbierenden Doppeltür abgedichtet worden. Während des Versuchs mussten beide Türen geschlossen sein.

Der Vorteil bei einer Darbietung der Stimuli im Freifeld gegenüber einer Darbietung über Kopfhörer liegt darin, dass bei der Verwendung von Kopfhörern Artefakte in der Lautheitswahrnehmung bedingt durch eine unterschiedliche Gehörgangsgröße der Probanden auftreten können. Bei Darbietung im Freifeld existiert dieses Problem nicht. Ein weiterer Vorteil der Freifeldmethode ist in ihrer besseren Reproduzierbarkeit zu sehen. Da dieser Versuch mit hörhilfetragenden Patienten fortgeführt werden soll, wäre es schwierig, den Kopfhörer stets genau über der externen Hörhilfe zu platzieren. Vergleichbarkeit und Reproduzierbarkeit wären somit eingeschränkt.

### 2.6.2 Probandenposition

Um Patienten des Klinikbetriebes sowie Probanden stets im exakt gleichen Abstand vom Lautsprecher positionieren zu können, war an der Decke über dem Sessel eine Schraubenmutter an einer Schnur befestigt. Der Abstand der Schraubenmutter zur Front des Lautsprechers betrug 1,0 m. An dieser Stelle waren auch die Schallpegelmessungen durchgeführt worden. Der Proband wurde mit dem Sessel so ausgerichtet, dass die Schraubenmutter knapp über der Kopfmittle (Schnittpunkt der Saggitalebene mit der Frontalebene durch die Gehörgänge) hing. So war gewährleistet, dass der Gehörgangsabstand vom Lautsprecher für alle Probanden 1,0 m betrug.

### 2.7 Instruktion der Probanden

Alle Probanden erhielten die gleichen Informationen. Es wurde darauf hingewiesen, dass die Reihenfolge der verrauschten Sätze zufällig war und alle Sätze mehrmals vorkommen konnten. Pro Testart existierten für jeden der 36 Sätze jeweils drei Varianten. Im Kategorientest kam jeder Satz bei sechs Listen (inkl. Übungsdurchlauf) also jeweils zweimal als Aussage, zweimal als Frage und zweimal als Ausruf vor. Analog wurde beim Sprechertest jeder Satz jeweils zweimal von den drei unterschiedlichen Sprechern gesprochen. So wurde keinem der Sätze bzw. der Satzvarianten der Vorzug gegeben und eine eventuelle Beeinflussung hierdurch vermieden. Um das Verhalten der Probanden während des Tests nicht zu beeinflussen, wurde ihnen nicht mitgeteilt, dass die Varianten der Sätze gleich oft vorkamen. Es lag also in dieser Hinsicht eine Verblindung der Versuchsteilnehmer vor. Andernfalls würde die Möglichkeit bestehen, dass sich die Probanden während des zweiten oder dritten Hörens eines Satzes an das bereits Gehörte erinnerten und sich so für eine ihrer Meinung noch nicht gehörte Variante entschieden.

Ebenso verhielt es sich mit den unterschiedlichen Antwortmöglichkeiten innerhalb einer Liste. Aus statistischen Gründen kamen in jeweils einer Liste für den Kategorientest je zwölf Aussagen, zwölf Fragen und zwölf Ausrufe vor. Ebenso waren in jeweils einer Liste des Sprechertests je zwölf Männer, zwölf Frauen und zwölf Kinder zu hören. Auf diese Weise blieb das Zufallsniveau der Antworten durch drei Antwortmöglichkeiten teilbar und die Wahrscheinlichkeit für eine korrekte Antwort durch reines Raten lag bei 33,3%. Wies man andererseits die Probanden auf diese Gleichverteilung hin, bestünde die Gefahr, dass beispielsweise nach zwölf zu hören geglaubten Männern automatisch in dieser Liste nur noch die Frau oder das Kind angegeben würden.

Umgekehrt bestand jedoch die Möglichkeit, dass die Probanden während des Tests die Gleichverteilung bemerkten und daraufhin ihr Verhalten änderten. Um dies möglichst auszuschließen, wurden die Teilnehmer auf absolute Zufälligkeit im Aufbau der Listen hingewiesen. Sie sollten sich ausschließlich auf das Gehörte konzentrieren.

### 2.7.1 Wortlaut der schriftlichen Probandeninstruktion

*„Der Zweck dieses Versuches ist es, Ihr prosodisches Verständnis der Sprache (mit Hauptaugenmerk auf die Intonation) bei unterschiedlichen Hörkonditionen, d. h. unterschiedlich lauten Störgeräuschen in Form eines Rauschens, festzustellen. Dabei befinden sich die verwendeten Schallpegel stets unter 90 dB, so dass eine Hörschädigung ausgeschlossen ist.*

*Vor der Durchführung des Tests wird ein Hörtest gemacht. Damit wird festgestellt, ob Sie als Proband tatsächlich „normalhörend“ und somit für den Test geeignet sind.*

*Der eigentliche Test gliedert sich in zwei Untertests. In Test 1 geht es um die Erkennung von Satzkategorien (Aussage, Frage, Ausruf), in Test 2 geht es um die Erkennung des Sprechers (Mann, Frau, Kind). Sowohl bei Test 1 als auch bei Test 2 hören Sie verschiedene Sätze mit verschiedenen lauten Störgeräuschen.*

*In Test 1 (Kategorientest) gibt es drei verschiedene Kategorien von Sätzen, und zwar Aussage, Frage und Ausruf: Aussagen sind emotionslose Sätze. Fragen können neutral oder emotional gefärbt (z. B. überrascht) ausgesprochen sein. Ausrufe sind emotional gefärbt, z. B. ärgerlich oder erfreut.*

*Bei allen drei Kategorien ist die Satzstruktur die gleiche. Sie sollen entscheiden, um welche Kategorie es sich handelt.*

*Bsp.: Aussage: Die Vorräte sind alle. Frage : Die Vorräte sind alle? Ausruf : Die Vorräte sind alle!*

*In Test 2 (Sprechertest) gibt es drei verschiedene Sprecher, und zwar Mann, Frau und Kind. Der Satzinhalt ist sekundär. Sie sollen nur feststellen, ob der Mann, die Frau oder das Kind gesprochen hat.*

*In beiden Tests kommen die beigemischten Störgeräusche in fünf verschiedenen Hörkonditionen vor, sind also unterschiedlich laut. Ihre Aufgabe ist es nun, trotz der*

## 2 Versuchsplanung und Methodik

*Störgeräusche herauszufinden, um welche Satzkategorie bzw. um welchen Sprecher es sich jeweils handelt. Dies wird Ihnen — abhängig vom Schallpegel der Störgeräusche — unterschiedlich schwer fallen.*

*Vor dem eigentlichen Test findet ein leichter Übungsdurchlauf statt. Die Störgeräusche sind hier zunächst ausschließlich leise, so dass Sie sich mit der Art des Tests vertraut machen können. Im Übungsdurchlauf ist der Satz lauter als das Rauschen, so dass Sie als Normalhörender den Satz mühelos verstehen.*

*Die Sätze werden Ihnen von einer CD vorgespielt und Ihre Antworten (Aussage, Frage, Ausruf bzw. Mann, Frau, Kind) von mir in vorbereitete Listen eingetragen. Für beide Tests gibt es jeweils fünf Listen mit jeweils 36 Sätzen. Jede Liste wird Ihnen in einer anderen Hörkondition vorgespielt, ist also mal einfacher und mal schwieriger zu verstehen. Alle Sätze können mehrmals vorkommen und ihre Reihenfolge ist zufällig.*

*Aus statistischen Gründen gibt es kein „Unentschieden“ oder „Weiß nicht“ und auch kein wiederholtes Anhören. Sie müssen sich immer für eine der drei möglichen Antworten entscheiden. Im Zweifel raten Sie bitte.*

*Versuchen Sie bitte, während des gesamten Tests möglichst konzentriert zu sein. Pausen zwischen den Einzeltests sind jederzeit möglich. Lassen Sie sich nicht irritieren, wenn Sie mehrere Male hintereinander die gleiche Kategorie oder den gleichen Sprecher hören, da die Reihenfolge absolut zufällig ist.*

*Sie können den Test jederzeit abbrechen, ohne dass Ihnen daraus irgendwelche Konsequenzen entstehen. Die bereits erhobenen Daten werden dann jedoch unbrauchbar und können nicht in die Arbeit mit einbezogen werden.*

*Die Daten werden zu wissenschaftlichen Zwecken erhoben und nur in anonymisierter Form ausgewertet. Die Daten werden verschlüsselt und wir verpflichten uns zur Einhaltung der datenschutzrechtlichen Bestimmungen.*

*Für Ihre Teilnahme an diesem Test und ihre Bemühungen möchte ich Ihnen schon mal im Voraus danken.*

*Ich habe obigen Text gelesen und zur Kenntnis genommen und erkläre meine Einwilligung zur Teilnahme an dem Versuch.*

*Datum, Teilnehmer/in“*

## 2.8 Versuchsablauf

Zuerst wurde das Tonaudiogramm in Kabine B aufgezeichnet. Nach bestandenerm Hörtest wurden die Probanden instruiert und der eigentliche Versuch in Kabine C durchgeführt. Dazu nahmen die Probanden im korrekt positionierten Sessel Platz und hörten zunächst einige Sätze der jeweils folgenden Testart als leichte Übungssequenz. Die SNR-Werte entsprachen für diesen Übungsdurchlauf +6 dB SNR für die Hauptgruppe. Nachdem etwaige Verständnisfragen geklärt waren, begann der Versuch. Die CDs mit den verrauschten Sätzen wurden über einen Computer der Marke Maxdata (Westra-CAD03-Software) mit Hilfe einer am Audiometer (Westra-Audiometer CAD03) befestigten HSM-Satztestkarte abgespielt. Das Audiometer wurde auf Freifelddarbietung und nominell 91 dB (Hauptgruppe) geschaltet. Bei dem zentral vor dem Probanden positionierten Lautsprecher handelte es sich um eine Audiometriebox LAB-1001 der Firma Westra Electronic GmbH. Anschließend wurden die Antworten der Probanden (Aussage, Frage, Ausruf bzw. Mann, Frau, Kind) manuell in vorbereitete Listen eingetragen. Nach jeder Testart wurden die Probanden im Hinblick auf Eigenbeobachtungen sowie Durchführungsschwierigkeiten befragt.

## 2.9 Methodik der statistischen Auswertung

In Vor- und Hauptversuch wurden zahlreiche Einzeldaten von insgesamt 36 Probanden erhoben. Diese Rohdaten wurden zusammengefasst und statistisch ausgewertet. Hierzu diente das Statistikprogramm *R* ([R Development Core Team, 2003](#)). Die Daten wurden nach unterschiedlichen Gesichtspunkten statistisch ausgewertet:

1. Leistungsspektrum der Probanden
2. Vergleich der Testarten
3. Streuung der Antworten
4. Urteilsverteilung in Bezug auf die dargebotenen Sprecher und Kategorien
5. Verstehbarkeit der einzelnen Sätze

Da mit diesem Test die Prosodieperzeption in fünf verschiedenen Schwierigkeitsstufen (−24 dB SNR, −18 dB SNR, −12 dB SNR, −6 dB SNR, 0 dB SNR) untersucht wurde, besteht eine Abhängigkeit von mehreren Einflussfaktoren. Diese wurden mit Hilfe von nichtlinearen logistischen Regressionen dargestellt. Es kamen jeweils drei

verschiedene Variationen der logistischen Regressionen in der Auswertung zum Einsatz. Hierbei wurden die ersten drei Gesichtspunkte gemeinsam statistisch dargestellt und unter dem Begriff „Darstellung der Gesamtdaten“ zusammengefasst.

### 2.9.1 Darstellung der Gesamtdaten

Die logistische Regression sollte das gesamte Leistungsspektrum der Probanden von der leichtesten Hörkondition (0 dB SNR) bis zur schwersten (-24 dB SNR) abbilden. Zur Darstellung der Abhängigkeit der korrekten Probandenantworten von den fünf Hörkonditionen als Einflussfaktoren wurden die SNR-Werte auf der x-Achse und die korrekten Antworten auf der y-Achse der Graphik aufgetragen. Zudem wurde die Regression durch eine asymptotische Annäherung an die Rate-Leistung und die maximale Leistung der Probanden dargestellt. Die beiden Grenzwerte  $a$  (Rate-Leistung) und  $b$  (maximale Leistung) waren somit wichtige Parameter bei der Erzeugung des Graphen. Die Beispielfunktion 2.11 veranschaulicht die Erzeugung der logistischen Regressionen. Folgende Formel beschreibt diese logistische Funktion:

$$L = f(a, b, s, SRT, SNR) = a + (b - a) \cdot \frac{1}{1 + e^{4s(SRT - SNR)}}$$

Dabei wird die Funktion  $f$  durch die vier Kurvenparameter  $a$ ,  $b$ ,  $s$  und  $SRT$  sowie den auf der x-Achse aufgetragenen  $SNR$  definiert. Der Parameter  $s$  kennzeichnet die Steigung und der  $SRT$  (speech reception threshold) den y-Achsenabschnitt, bei dem 50 % der Sprache wahrgenommen wird und somit der Wendepunkt der Kurve liegt.

Durch die asymptotische Annäherung der Kurve an die Rate-Leistung (0,33) und die maximale Leistung (1,0) ist das gesamte Leistungsspektrum der Probanden erfassbar und der Bereich, in dem die Kurve liegt, nach oben und unten begrenzt. Der Wendepunkt wird durch den  $SRT$  definiert und markiert den Übergang zwischen leichteren Bereichen mit vielen richtigen Antworten und schwereren Bereichen. Die Lage des Wendepunktes ermöglicht somit Aussagen über den SNR-Wert, ab dem die Versuchsteilnehmer zunehmend Schwierigkeiten bei der Prosodieperzeption hatten und vermehrt falsche Antworten gaben. Die Steigung im Wendepunkt (Steilheit der Funktion) zeigt an, wie steil oder flach der Übergang ausgeprägt ist. Ein flacher Übergang würde eine langsame Abnahme der Leistung bedeuten, während ein steiler Übergang einen schnellen Leistungsverlust ab einem bestimmten SNR-Wert kennzeichnen würde. Ein Vergleich der Testarten wurde durch eine Abbildung



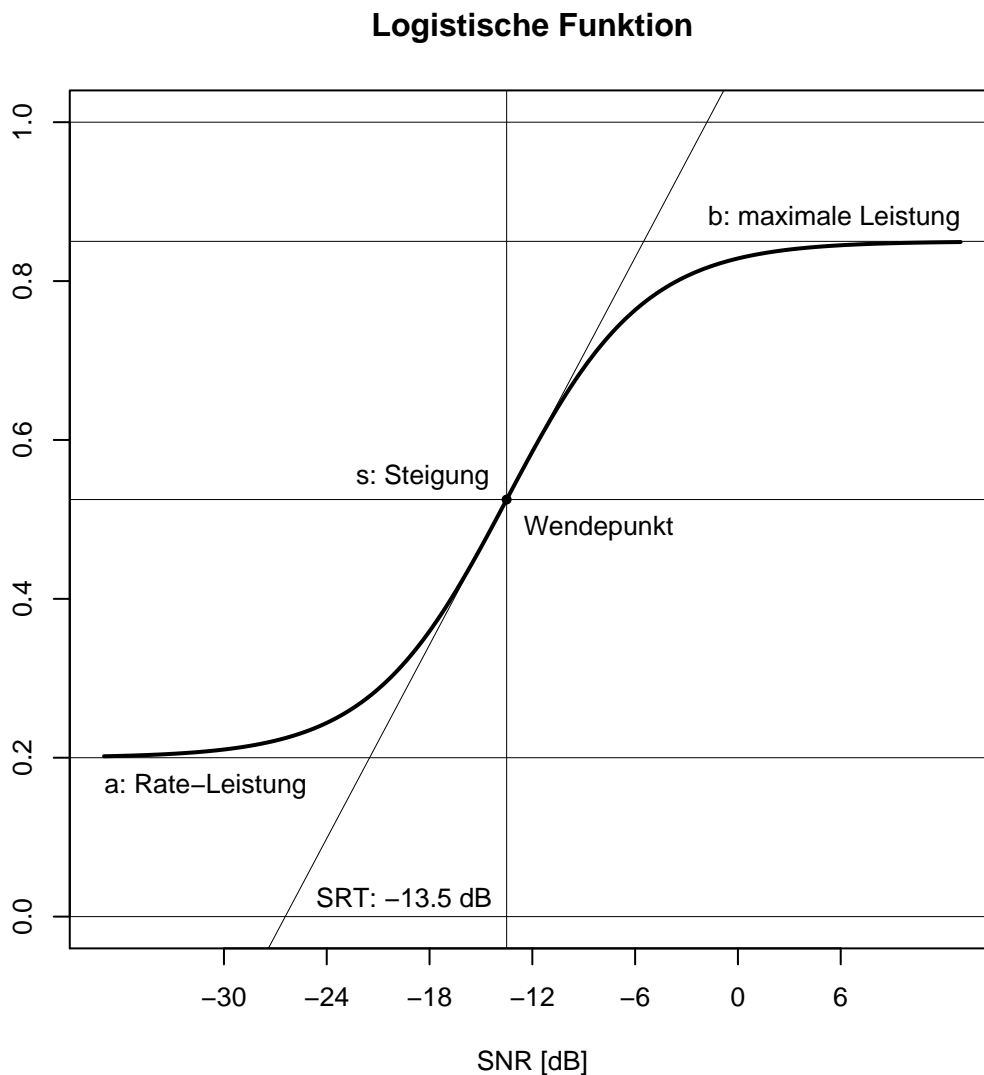


Abbildung 2.11: Beispiel einer logistischen Funktion

der Regressionen für Sprecher- und Kategorientest in derselben Graphik ermöglicht. Die jeweiligen Lagen der Wendepunkte würden Aussagen bezüglich der Schwierigkeit der Testarten ermöglichen. Eine Darstellung der Einzelkurven der Probanden soll weiterhin die Streuung der Kurven und somit die Varianz der Ergebnisse verdeutlichen.

Bei der Erzeugung der Funktionen kamen zwei unterschiedliche Methoden zur Anwendung:

- Die kompletten Daten aller Probanden lagen in Form einer großen Punktwolke vor, durch die jeweils eine Kurve für die jeweilige Testart gelegt wurde. Begrenzungsparameter waren  $a$  (0,33) und  $b$  (1,0).

- Als Kurvenparameter wurden die Mittelwerte der Einzelkurven der Probanden genommen. Die Funktion wurde somit zusätzlich zu den Begrenzungsparametern  $a$  und  $b$  durch die Parameter Steigung  $s$  und SRT definiert.

Trotz der Anwendung unterschiedlicher Methoden sind die Kurvenverläufe nahezu deckungsgleich, so dass von ihrer beider Eignung zur Darstellung der Daten ausgegangen werden kann (s. die jeweils breiter gezeichneten Funktionen der Abbildungen 3.2 und 3.3). Da bei einer Erzeugung der Kurve mit Hilfe der Punktwolke beispielsweise steile Übergänge im schweren Bereich mit steilen Übergängen im leichten Bereich kombiniert wurden, verläuft diese etwas flacher (in den Abbildungen 3.2 und 3.3 jeweils gestrichelt dargestellt).

### 2.9.2 Urteilsverteilung

Um die Verteilung der Urteile, mögliche Bevorzugungen oder sogar Verwechslungen der Sprecher bzw. Kategorien darstellen zu können, wurden zunächst alle Daten in Form von Kontingenztabellen sowie einem Blasendiagramm abgebildet. Im nächsten Schritt wurden die Probandenurteile sowohl in Bezug auf die jeweils dargebotenen Sprecher M, W oder K als auch in Bezug auf die Kategorien A, F oder R graphisch dargestellt.

Zur Anwendung kamen hierbei ebenfalls nichtlineare logistische Regressionen. Diese waren zunächst durch die Kurvenparameter Steigung  $s$  sowie den SRT definiert. Durch die gleichzeitige Abbildung dreier Graphen entsprechend den drei Antwortmöglichkeiten für die jeweils dargebotenen Sprecher bzw. die jeweiligen Kategorien, mussten alle Urteile für einen SNR in ihrer Summe 1 ergeben. Wurden demzufolge beispielsweise bei 0 db SNR 95 % der Urteile für den männlichen Sprecher abgegeben, mussten sich die restlichen 5 % der Urteile auf Frau und Kind verteilen.

Zusätzlich wurde eine asymptotische Annäherung der Funktion an die 1,0-Marke der maximalen Leistung vorgegeben. Die untere Begrenzung war im Gegensatz zur Funktion der Darstellung der Gesamtdaten mit 0,0 definiert, da Aussagen über die Verteilung der Urteile getroffen werden sollten.

Ferner gab es für jede Antwortmöglichkeit eine Ratewahrscheinlichkeit. Dies beruht auf folgendem Umstand: Waren die Probanden auf reines Raten angewiesen, gaben sie die Urteile nicht mit einer Gleichverteilung von jeweils 33,3 % an. Bestimmte Antwortmöglichkeiten wurden anderen vorgezogen. So gab es für jede Antwortmöglichkeit eine andere Rateleistung  $a$ . Um die Funktionen noch besser zu definieren, wurden sie durch diese Werte geplottet und dadurch im Wert  $a$  gekoppelt. Die

maximale Leistung  $b$  entspricht der Summation aller Urteile für den jeweiligen SNR auf 1 und somit der oberen Begrenzung.

### 2.9.3 Verstehbarkeit der Sätze

Zur Darstellung der Verstehbarkeit der Sätze wurden die Urteilsverteilungsfunktionen für jeden der insgesamt 36 Sätze aufgesplittet. Da aber somit wesentlich weniger Daten für jeden Satz verfügbar waren und diese eine zum Teil erhebliche Streuung aufwiesen, wurden die Punkte zur Veranschaulichung lediglich durch Striche untereinander verbunden. In den ersten drei Spalten sind die Urteile satzweise für die dargebotenen Sprecher und Kategorien dargestellt. Unterlegt wurden diese Graphiken mit den Funktionen der Urteilsverteilung aller Sätze (vgl. Abbildungen 3.5–3.10).

Die vierte Spalte stellt alle korrekten Urteile für jeden Satz dar. Die unterlegte Kurve entspricht der logistischen Funktion der Gesamtdaten der jeweiligen Testart in Abbildung 3.1 und gibt somit die korrekten Urteile für alle Sätze einer Testart an.



## 3 Ergebnisse

### 3.1 Darstellung der Gesamtdaten

In diesem Abschnitt werden zunächst die Ergebnisse der Probanden für die beiden Testarten Sprecher und Kategorie anhand Abbildung 3.1 erläutert. Auf der Abszisse sind die SNRs von  $-30$  dB bis  $+6$  dB aufgeführt. Die Ordinate zeigt die Resultate der Probanden in einem Bereich von  $0,33$  als Grenzwert der Ratewahrscheinlichkeit bis  $1,0$  als obere Grenze des fehlerfreien Antwortens. Die blauen Punkte sowie die blaue Kurve kennzeichnen die Resultate aller Versuchspersonen für den jeweils geprüften SNR der Testart Sprecher. Orange gezeichnete Werte stehen für die Testart Kategorie. Dabei enthält jeweils ein Punkt bei  $5400$  Antworten pro Testart  $1080$  Einzelantworten.

Definitionsgemäß weisen beide Funktionen eine asymptotische Annäherung an die Grenzbereiche der Rate-Leistung und der maximalen Leistung auf (s. Abschnitt 2.9), d. h. die Parameter der logistischen Funktion  $f$  wurden zu  $a = 0,33$  und  $b = 1,0$  vorab festgelegt. Die relativen Häufigkeiten weisen naturgemäß Abweichungen auf: Bei  $-24$  dB SNR gaben die Probanden im Sprechertest mit knapp  $50\%$  immer noch viele richtige Antworten und sind damit relativ weit von der erwarteten Ratewahrscheinlichkeit  $33,3\%$  entfernt. Im Kategorientest hingegen liegen die Probanden bei  $-24$  dB SNR relativ nahe an der  $33,3\%$ -Grenze.

Auffallend ist, dass die Probanden im Kategorientest bei  $0$  dB SNR nicht die obere Marke ( $100\%$ ) erreichen, sondern auch in den leichteren Hörkonditionen schlechtere Resultate als im Sprechertest aufweisen. Bei  $0$  dB SNR gaben die Versuchsteilnehmer im Sprechertest annähernd zu  $100\%$  richtige Antworten.

Bei Betrachtung der Lage der Wendepunkte fällt auf, dass der Wendepunkt des Sprechergraphen links des Wendepunkts des Kategoriengraphen liegt. Mit einem Wert von  $-14,72$  dB SNR für den Wendepunkt des Sprechertests besteht ein deutlicher Unterschied zu dem des Kategorientests mit  $-11,65$  dB SNR.

Des Weiteren weist der Sprechergraph eine geringere Steigung, d. h. einen flacheren Übergang zwischen leichtem und schwerem Hörbereich, auf. Verfolgt man den Verlauf der Kurven für beide Testarten in den Bereich der schwereren Hörkondi-

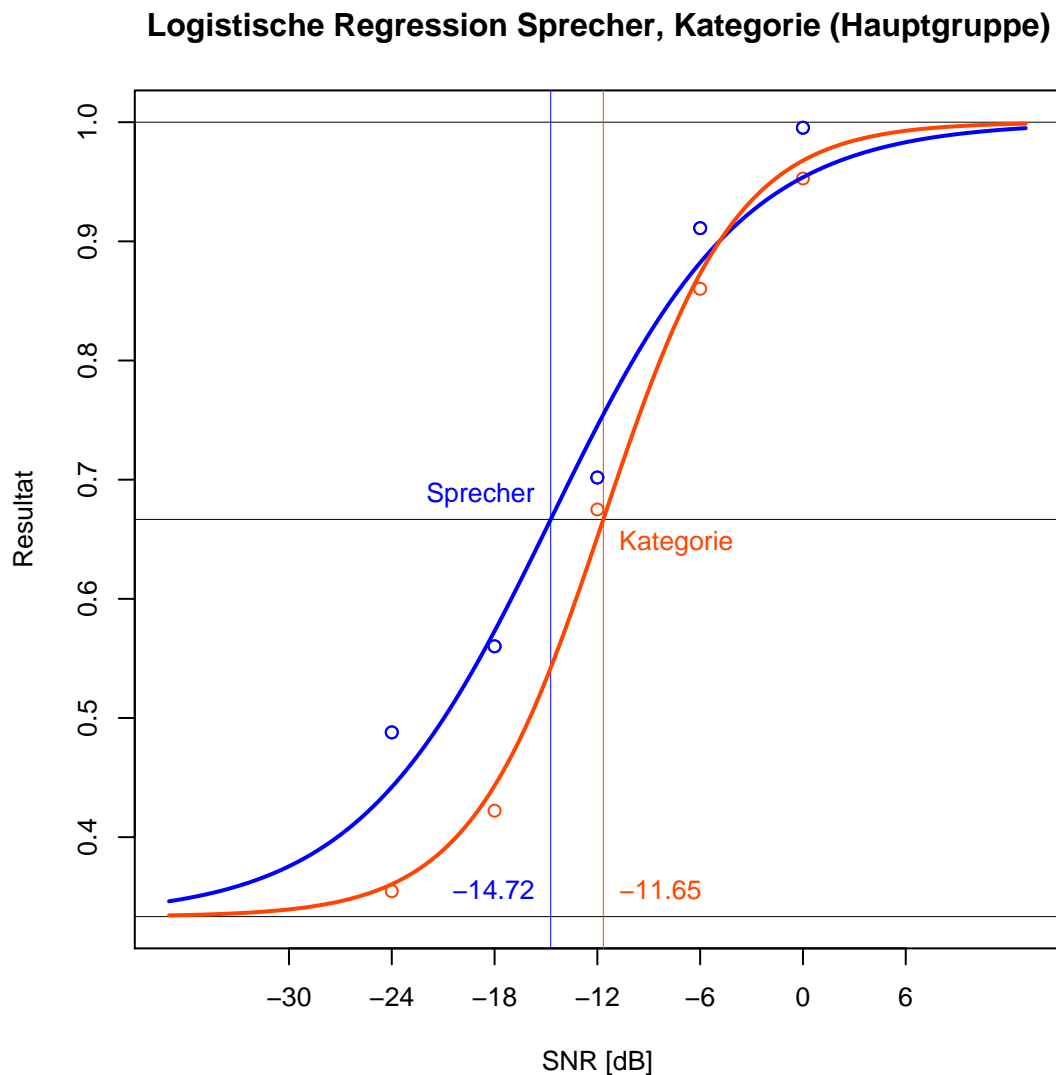


Abbildung 3.1: Logistische Regressionen für die Testarten Sprecher und Kategorie

tionen, vergrößert sich der Abstand zwischen beiden Funktionen. Der Sprechertest erzielt also etwas bessere Antworten bei den leichteren und bedeutend bessere bei den schwereren Hörkonditionen.

Die Abbildungen 3.2 und 3.3 zeigen die logistischen Regressionen für die Testarten Sprecher und Kategorie mit ihren jeweiligen Einzelkurven der Probanden, die als dünne helle Linien abgebildet sind. Die breiter gezeichneten dunkleren Kurven stellen die Daten aller Probanden zusammengefasst dar: Die jeweils durchgezogene Kurve wurde mit der „Punktwolke“ erzeugt und die jeweils gestrichelte mit den Mittelwerten der Einzelkurven der Probanden (vgl. Abschnitt 2.9.1).

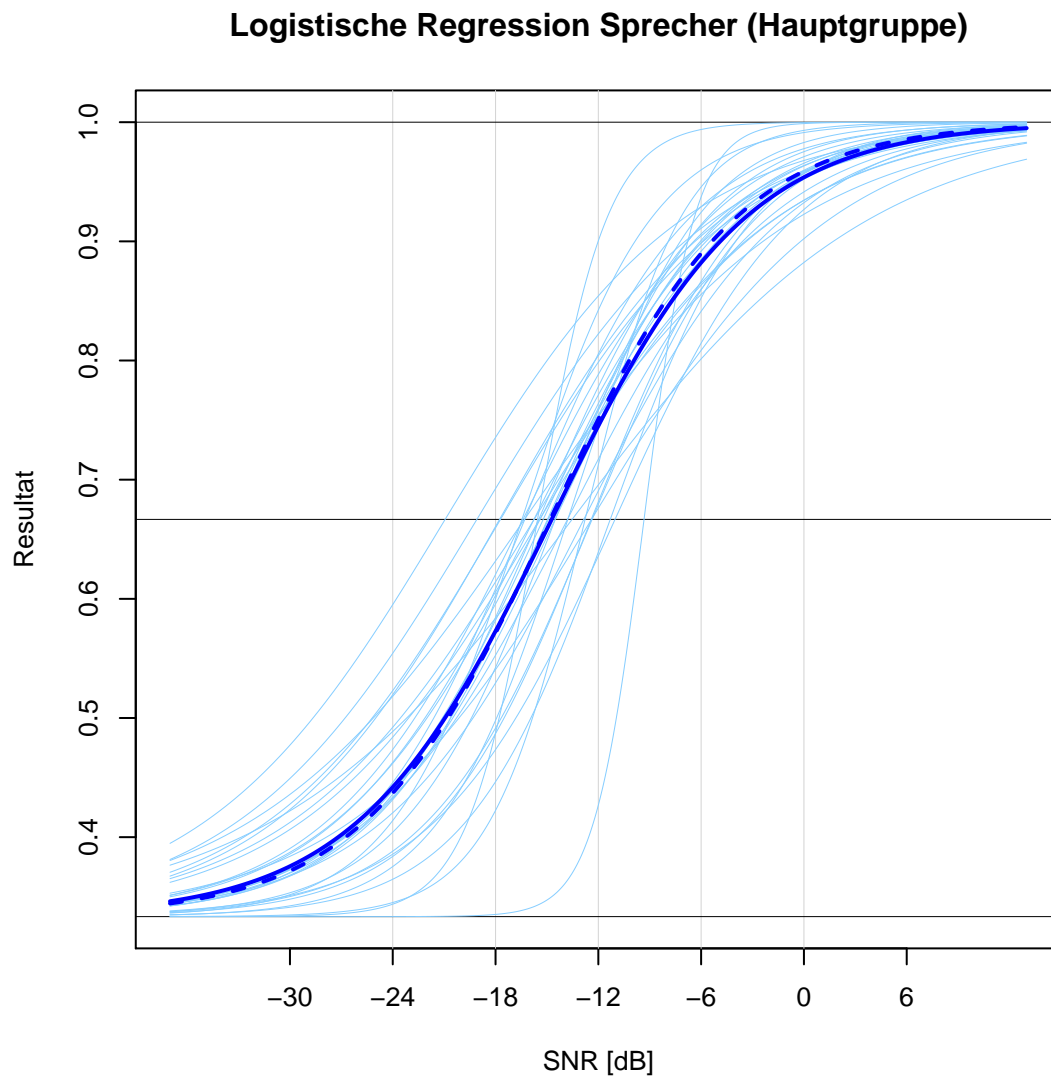


Abbildung 3.2: Logistische Regressionen für die Testart Sprecher inkl. der Einzelkurven der Probanden

Bei Betrachtung der Abbildung 3.3 ist auffallend, dass die Einzelkurven eine relativ geringe Streuung aufweisen. Die Wendepunkte aller Kurven liegen zwischen ca.  $-8,5$  dB SNR und ca.  $-14$  dB SNR.

Die Einzelkurven der Probanden für die Testart Sprecher in Abbildung 3.2 weisen eine größere Streuung auf. Ihre Wendepunkte liegen in einem Bereich von ca.  $-9$  dB SNR bis ca.  $-21$  dB SNR.

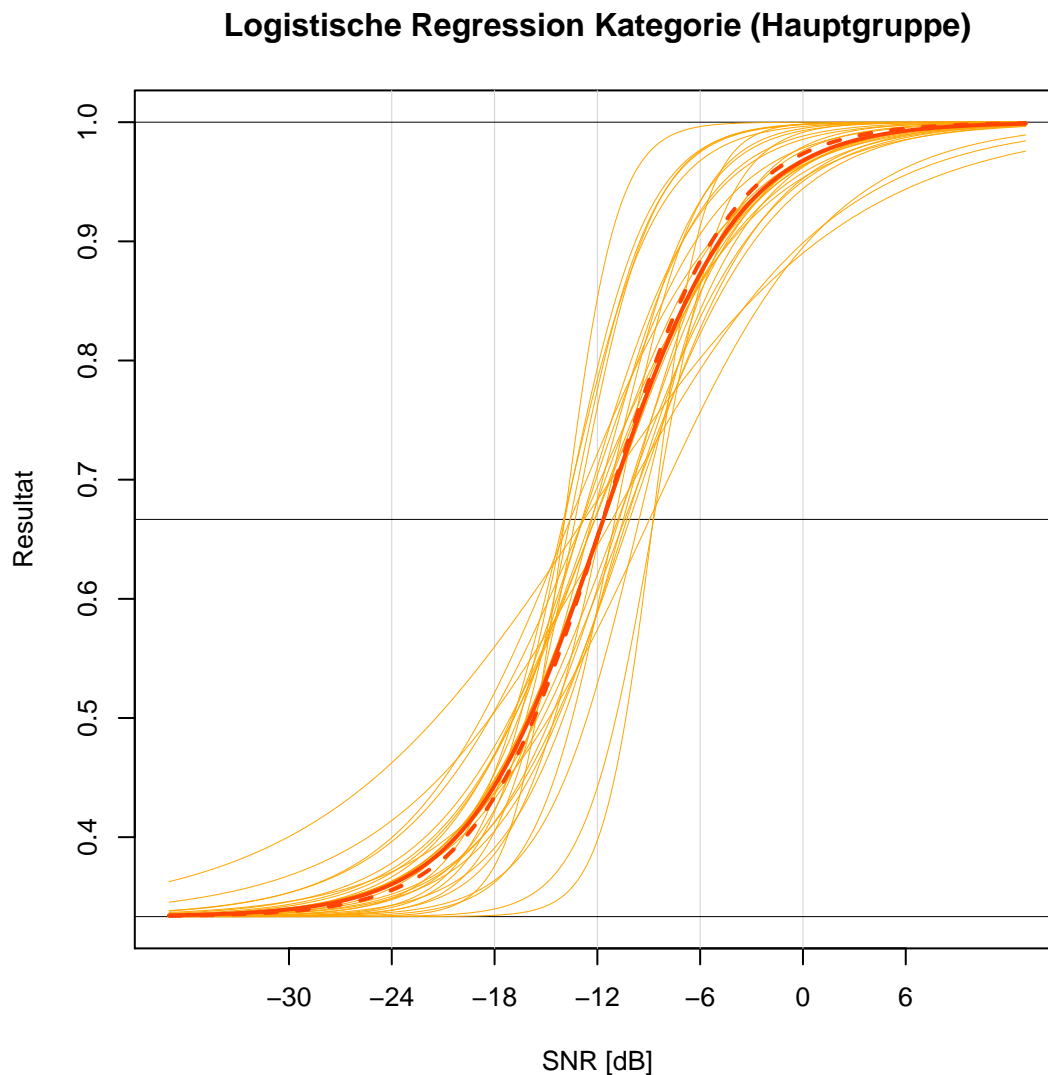


Abbildung 3.3: Logistische Regressionen für die Testart Kategorie inkl. der Einzelkurven der Probanden

### 3.2 Urteilsverteilung

Nachdem nun ein Gesamtüberblick über die Leistungen der Probanden vorliegt, soll näher betrachtet werden, wie die abgegebenen Urteile auf die dargebotenen Sprecher bzw. Kategorien verteilt sind. Im folgenden sind die Probandenurteile in Abhängigkeit der SNRs in einem Blasendiagramm (s. Abbildung 3.4) dargestellt. Die linke Spalte zeigt die Urteilsverteilungen für den Sprechertest, rechts sind die Probandenurteile der Testart Kategorie dargestellt. Zuerst sind die Graphiken der leichtesten



Hörkondition bei 0 dB SNR dargestellt und im weiteren Verlauf bis hin zur schwierigsten Hörbedingung bei  $-24$  dB SNR aufgezeichnet. Auf der Abszisse sind die erkannten Sprecher M, W, K und die erkannten Kategorien A, F und R aufgetragen. Die Ordinate zeigt die jeweils dargebotenen Sprecher (links) bzw. Kategorien (rechts). Eine Blase ist umso größer, je mehr richtige Antworten in ihr vereint werden. Die genauen Zahlenwerte sind den zugehörigen Kontingenztabellen im Anhang B zu entnehmen.

### 3.2.1 Urteilserkennung

Zur Beantwortung der Frage, wie gut bzw. mit wie vielen richtigen Antworten der jeweilige Sprecher oder die jeweilige Kategorie erkannt wurde, ist eine Betrachtung der Diagonalen von links unten nach rechts oben in den Blasendiagrammen erforderlich. Je größer eine Blase gezeichnet ist, umso mehr korrekte Urteile sind für den jeweiligen Sprecher bzw. die jeweilige Kategorie in ihr vereint. Im Optimalfall der fehlerfreien Urteilsabgabe wären in der Diagonalen nur jeweils drei große Blasen zu sehen, während die restlichen Felder leer blieben. Da pro SNR und Testart je 1080 Urteile abgegeben wurden, entspräche eine solche Blase 360 Urteilen.

Bei den folgenden Erläuterungen des Blasendiagramms soll aus Gründen der Übersichtlichkeit stets zuerst auf den Sprechertest und im Anschluss auf den Kategorientest eingegangen werden.

Im Sprechertest liegen bei 0 dB SNR für alle Sprecher sehr viele richtige Antworten vor, die Urteile sind fast diagonal verteilt. Mann und Kind wurden stets korrekt zugeordnet. Im weiteren SNR-abhängigen Verlauf wurden alle Sprecher kontinuierlich schlechter erkannt. Im Blasendiagramm fällt die Abnahme der richtigen Antworten für den Mann am geringsten aus, während die Erkennung des kindlichen Sprechers ab  $-6$  dB SNR bis hin zu  $-24$  dB SNR stark nachlässt. Auffällig ist die große Diskrepanz der richtigen Antworten bei  $-24$  dB SNR für die jeweiligen Sprecher. Der Mann weist um das Vierfache mehr richtige Antworten auf als das Kind.

Die Graphik für den Kategorientest bei 0 dB SNR sowie die Kontingenztafel B.8 im Anhang zeigen, dass die Frage am besten erkannt wurde. Im SNR-abhängigen Verlauf bis  $-12$  dB SNR wurde der Ruf in Relation zu Aussage und Frage zunehmend schlechter erkannt, ab  $-12$  dB SNR fällt die Erkennung des Rufs etwas weniger stark ab. Die Frage wurde ab  $-6$  dB SNR im weiteren Verlauf kontinuierlich schlechter erkannt, bei  $-24$  dB SNR wurde sie am seltensten gewählt. Im Bereich der schwierigeren Hörkonditionen wurde die Aussage am häufigsten erkannt.

### 3 Ergebnisse

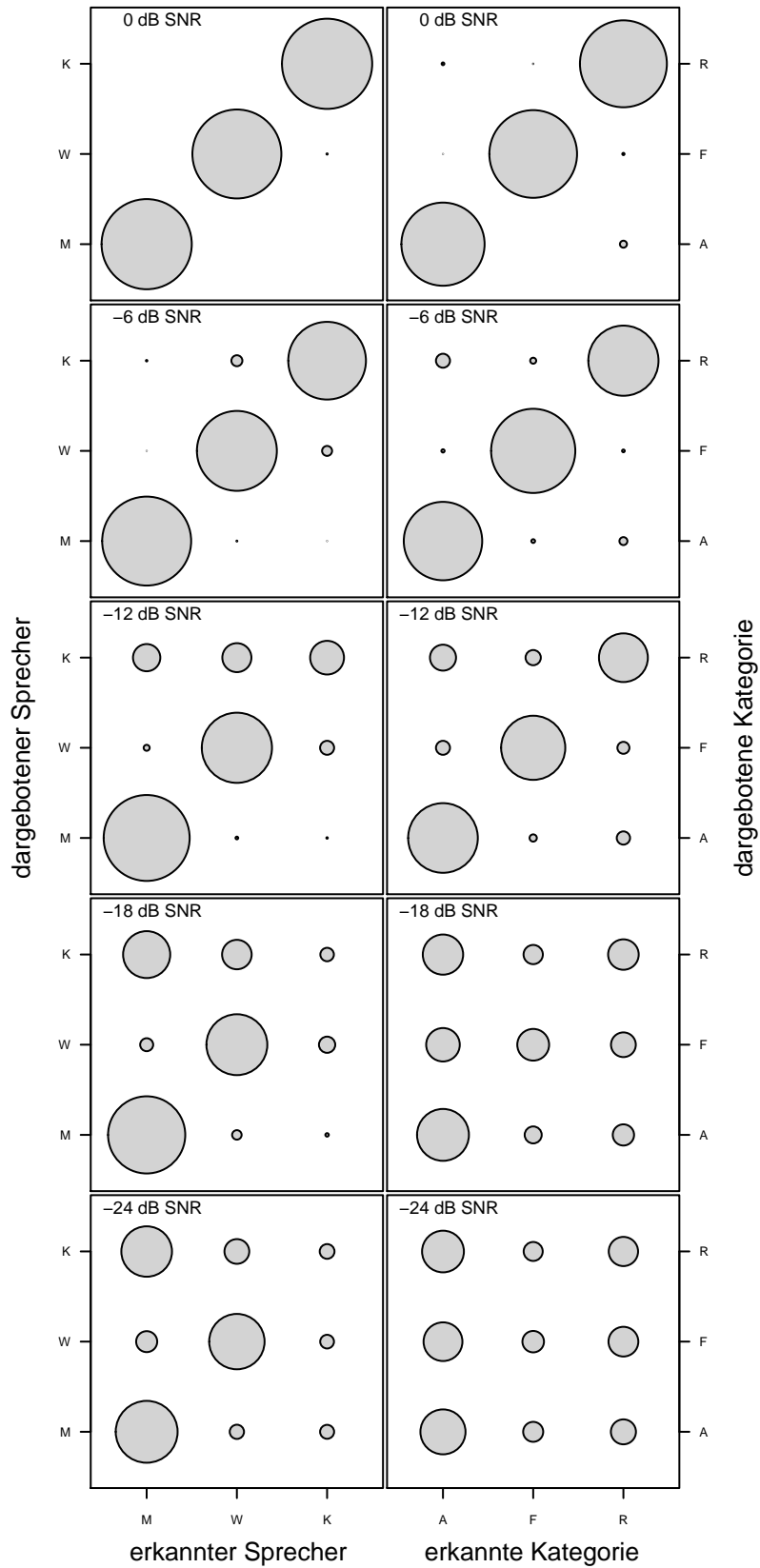


Abbildung 3.4: Urteilsverteilung der Sprecher und Kategorien in Abhängigkeit vom SNR

Im Vergleich der beiden Testarten ist die Anzahl der korrekten Urteile bei den Sprechern (mit Ausnahme des kindlichen Sprechers) bis hin zu  $-24$  dB SNR stets höher als bei den Kategorien.

### 3.2.2 Verwechslungen

Im folgenden soll die Art der Verwechslungen erläutert werden. Es soll untersucht werden, wie die falschen Antworten verteilt sind und ob es „umgekehrte Treffer“ gibt (s. hierzu Abbildung 3.4). Bei einem „umgekehrten Treffer“ erfolgte beispielsweise die Abgabe des Urteils Mann bei Darbietung der Frau bei gleichzeitiger Abgabe des Urteils Frau bei Darbietung des Mannes. Hierzu werden die Diagramme jeweils zeilenweise betrachtet.

Die Graphiken der Testart Sprecher zeigen, dass der Mann und das Kind bei 0 dB SNR eindeutig erkannt wurden. Im weiteren Verlauf wurde der Mann öfter mit der Frau verwechselt. Ab  $-18$  dB SNR kann jedoch ebenfalls eine Zunahme der Verwechslungen mit dem Kind beobachtet werden. Die Frau wurde als einzige der Sprecher bereits bei 0 dB SNR fünfmal mit dem Kind verwechselt (s. Tabelle B.2). Während unter den leichteren Hörbedingungen Verwechslungen der Frau mit dem Kind überwiegen, ist im schwierigeren Bereich ein relativer Anstieg der Verwechslungen mit dem Mann zu verzeichnen. Bei Darbietung des Kindes erfolgte bei  $-6$  dB und  $-12$  dB SNR der Großteil der Verwechslungen mit der Frau. Hier liegen „umgekehrte Treffer“ vor. Mit zunehmender Lautheit des Rauschens steigen jedoch bei Darbietung des Kindes die Urteile für den Mann an, bis diese bei  $-18$  dB SNR die Urteile für die Frau und sogar das Kind selbst übertreffen. Da bei  $-24$  dB SNR bei Darbietung des Mannes der Anteil der Fehlantworten für das Kind ansteigt, während bei Darbietung des Kindes mit der Mehrzahl der Urteile der Mann gewählt wurde, scheint es auch hier „umgekehrte Treffer“ zu geben.

Im Kategorientest liegen bei den beiden leichtesten Hörbedingungen mehr Verwechslungen vor als im Sprechertest. Für die Aussage existieren bei 0 dB SNR 28 Verwechslungen mit dem Ruf (s. Tabelle B.8), ab  $-6$  dB SNR kommen Verwechslungen mit der Frage hinzu. Im weiteren Verlauf überwiegen ebenfalls stets Verwechslungen mit dem Ruf. Der Ruf wurde bei 0 dB SNR hauptsächlich mit der Aussage verwechselt. Hierbei handelt es sich um „umgekehrte Treffer“, da bei Darbietung der Aussage der Ruf als alleinige Falschantwort angegeben wurde. Ab  $-6$  dB SNR nehmen die Verwechslungen des Rufes mit der Aussage bedeutend zu, bis die Aussage in der schwersten Hörkondition sogar mehr Urteile zu verzeichnen hat. Bei Darbie-

### 3 Ergebnisse

tung der Kategorie Frage erfolgten bei 0 dB SNR die meisten Verwechslungen mit dem Ruf. Mit zunehmender Schwierigkeit der Hörkonditionen steigt die Anzahl der Verwechslungen mit den übrigen Kategorien kontinuierlich an, bis Aussage und Ruf bei  $-24$  dB SNR sogar öfter gewählt wurden als die Frage.

#### 3.2.3 Urteilshäufigkeit

Es soll nun die Urteilshäufigkeit im SNR-abhängigen Verlauf geschildert werden. Die Urteilshäufigkeit gibt an, wie oft für ein Urteil – egal ob richtig oder falsch – gestimmt wurde. Hierzu sind die Blasen in den jeweiligen Spalten der erkannten Sprecher oder Kategorien in ihrer Summe zu betrachten (s. Abbildung 3.4). Die zugehörigen Zahlenwerte sind in der jeweils letzten Zeile der Kontingenztabelle abgebildet (s. Anhang B).

Bei Betrachtung des Sprechertests hat der Mann im SNR-abhängigen Verlauf einen erheblichen Anstieg der Urteile zu verzeichnen. Die Anzahl der abgegebenen Urteile für die Frau ist vergleichsweise konstant, während das Kind durch eine stetige Urteilsabnahme auffällt. In den schwierigeren Hörkonditionen sind die Probandenurteile zum großen Teil auf Mann und Frau verteilt, wobei der Mann die meisten Urteile zu verzeichnen hat. Dem kindlichen Sprecher verbleiben somit nur sehr wenige Urteile.

Im Kategorientest steigt die Summe der Urteile für die Aussage im SNR-abhängigen Verlauf kontinuierlich an. Die Summe der Urteile für die Frage ist bis ca.  $-12$  dB SNR relativ konstant und ist dann durch eine starke Abnahme gezeichnet. Dem Ruf wurde – bedingt durch Verwechslungen bei Darbietung der Aussage – bei 0 dB SNR der Vorzug gegeben. Im SNR-abhängigen Verlauf kommt es durch die zunehmende Beliebtheit des Urteils Aussage bei  $-12$  dB SNR zu einer relativen Urteilsabnahme. In der  $-24$  dB SNR-Graphik sind die Urteile eher auf die Randbereiche verteilt. Bei Darbietung aller Kategorien Aussage, Frage und Ruf entschieden sich die Probanden in den meisten Fällen für die Aussage. Der Ruf wurde als Antwort ebenfalls oft gewählt, während die Frage am unbeliebtesten war. Sie wurde nur knapp halb so oft wie die Aussage angegeben.

### 3.3 Ratewahrscheinlichkeiten

Zur Veranschaulichung sind die Probandenurteile hier auf eine andere Weise graphisch dargestellt. Die Erzeugung der Graphen (s. Abbildungen 3.5–3.10) wurde be-

Tabelle 3.1: Ratewahrscheinlichkeiten aller Sprecher und Kategorien in Prozent

|           |         |       |       |
|-----------|---------|-------|-------|
| Sprecher  | Mann    | Frau  | Kind  |
| Prozent   | 50,3    | 31,9  | 17,8  |
| Kategorie | Aussage | Frage | Ausuf |
| Prozent   | 45,7    | 23,1  | 31,2  |

reits in Abschnitt 2.9.2 besprochen. Für beide Testarten gibt es jeweils drei Abbildungen. In jeder Abbildung sind die Urteile der Probanden für die jeweils dargebotene Kategorie Aussage, Frage oder Ruf bzw. die Sprecher Mann, Frau oder Kind in Abhängigkeit vom SNR dargestellt. Die Abszisse gibt die SNRs der fünf Hörkonditionen an, während auf der Ordinate die Antworten der Probanden als relative Häufigkeit aufgetragen sind. Anders als bei den logistischen Funktionen in Abschnitt 3.1 ist hier eine asymptotische Annäherung an  $y = 0,0$  als untere Begrenzung sowie  $y = 1,0$  als obere Begrenzung vorgegeben, da Häufigkeit und Verteilung der abgegebenen Urteile dargestellt werden. Aufgrund der Verteilung der Urteile für die jeweils dargebotenen Sprecher bzw. die jeweiligen Kategorien auf je drei Antwortmöglichkeiten pro Testart ergibt die Summe der Urteile für jeden SNR stets 1 (vgl. Abschnitt 2.9.2).

Im Vergleich der drei Abbildungen pro Testart untereinander wird deutlich, dass sich die Kurven in ihrer Verlängerung über die  $-24$  dB SNR-Marke hinaus stets den gleichen Werten annähern. Dies stellt die Ratewahrscheinlichkeiten für Sprecher sowie Kategorien dar (vgl. Abschnitt 2.9.2). Beim Sprechertest erreicht der Mann ca. 50% der Urteile, für Frau und Kind wurden entsprechend weniger Urteile abgegeben. Im Kategorientest ist die Aussage am beliebtesten, gefolgt von Ruf und Frage. Die genauen Ratewahrscheinlichkeiten sind in Tabelle 3.1 dargestellt.

### 3 Ergebnisse

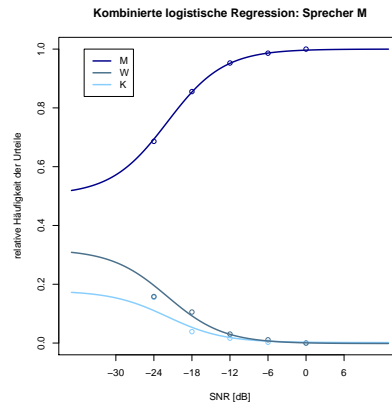


Abbildung 3.5: Urteilsverteilung bei Darbietung des Sprechers M in Abhängigkeit vom SNR

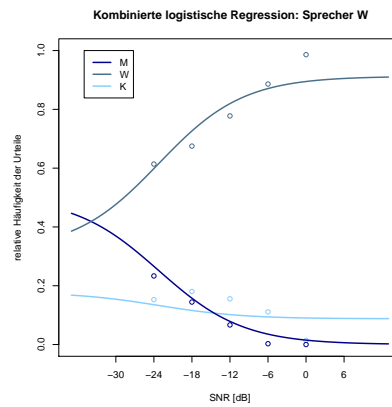


Abbildung 3.6: Urteilsverteilung bei Darbietung des Sprechers W in Abhängigkeit vom SNR

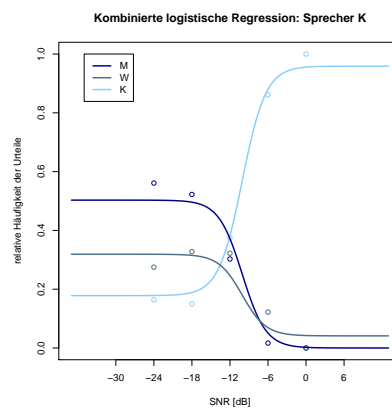


Abbildung 3.7: Urteilsverteilung bei Darbietung des Sprechers K in Abhängigkeit vom SNR

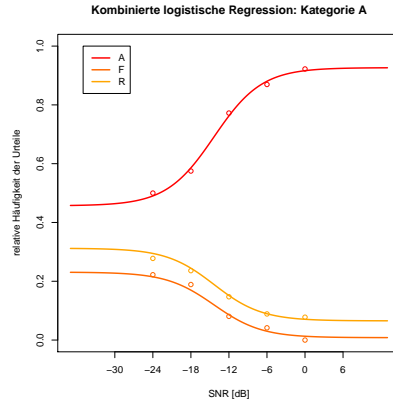


Abbildung 3.8: Urteilsverteilung bei Darbietung der Kategorie A in Abhängigkeit vom SNR

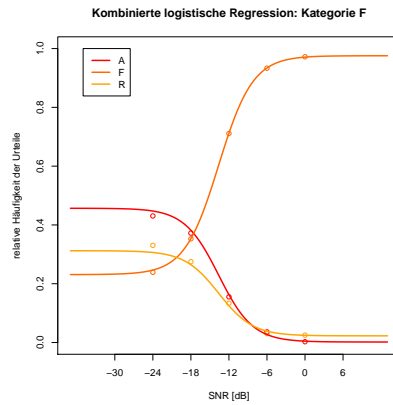


Abbildung 3.9: Urteilsverteilung bei Darbietung der Kategorie F in Abhängigkeit vom SNR

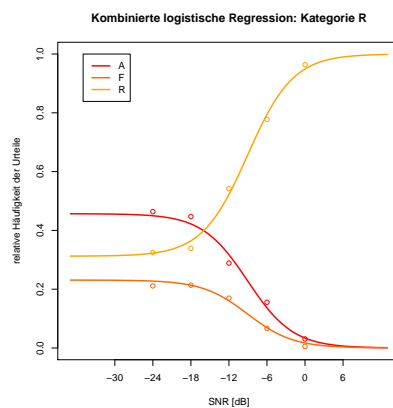


Abbildung 3.10: Urteilsverteilung bei Darbietung der Kategorie R in Abhängigkeit vom SNR

## 3.4 Verstehbarkeit der Sätze

Die Graphiken aller Sätze für beide Testarten sind im Anhang C dargestellt. Es sollte festgestellt werden, ob Auffälligkeiten oder Abweichungen einzelner Sätze bezüglich der Kategorien- bzw. Sprechererkennbarkeit vom Gesamtdatensatz existieren. Es sollten Aussagen bezüglich der Eignung der einzelnen Sätze getroffen werden. Daraufhin könnte nach Gründen und Erklärungen für die Ausreißersätze gesucht werden und diese durch neue verbesserte Sätze ersetzt werden. Es würde dann ein Testmaterial mit ausschließlich geeigneten Sätzen vorliegen, das ebenfalls in einem weiteren Schritt an Probanden getestet werden müsste.

Hier sollen einige exemplarische Sätze aufgezeigt werden, die größere Abweichungen von den Graphen der Gesamtheit aller Sätze aufweisen und durch ihre Diskrepanz Aufmerksamkeit auf sich ziehen.

### 3.4.1 Erkennbarkeit der Kategorien

Zunächst sollen die graphischen Darstellungen der Sätze für den Kategorientest näher betrachtet werden (s. Abbildungen C.1–C.6). Es sind einige auffällige Unterschiede sowohl in der Urteilsverteilung als auch in Bezug auf die Anzahl der korrekten Antworten erkennbar. Es werden vier Beispiele gezeigt:

Satz Nr. 3 („Schön haben wir es hier“) weist starke Verwechslungen der Kategorien Aussage und Ruf in den leichteren Hörkonditionen auf. So wurde bei Darbietung der Aussage diese bei 0 dB SNR nur zu 60 %, bei –6 dB SNR zu 70 % gewählt. Die übrigen Probanden glaubten einen Ruf zu hören. Bei Darbietung des Satzes als Ruf gaben bei 0 dB SNR nur 70 % und bei –6 dB SNR nur 50 % der Probanden ihr Urteil für die Kategorie Ruf ab, während die übrigen Urteile der Aussage zugeordnet wurden. Die Frage hingegen wurde bis in den schwierigen Bereich der Hörkonditionen gut erkannt. Diese Beobachtungen spiegeln sich in der vierten Spalte wider. Die Anzahl der korrekten Antworten in den beiden leichtesten Hörkonditionen liegt deutlich unter dem Gesamtdurchschnitt.

Bei Betrachtung des Satzes Nr. 21 („Der Maler ist wirklich begabt“) fällt eine starke Tendenz zum Urteil Ruf auf. Sowohl bei Darbietung des Satzes als Aussage als auch als Frage kam es zu einer Urteilshäufung für den Ruf. Die Graphik für die Kategorie Aussage verdeutlicht dies mit einer Überkreuzung der Kurven Aussage und Ruf zwischen 0 dB SNR und –6 dB SNR. Für die Aussage wurden zu Gunsten des Rufes nur 40 % der Urteile abgegeben. In der Graphik für die Kategorie Frage wird



in den mittelschweren Hörkonditionen eine Häufung des Urteils Ruf zu Lasten der Frage gezeigt. Die vierte Spalte zeigt relativ wenige richtige Urteile in den leichteren Hörbedingungen und verdeutlicht damit die genannten Beobachtungen.

Ebenso zeigt Satz Nr. 22 („Klettern kann gefährlich sein“) Fehlinterpretationen von Aussage und Ruf. Bei Darbietung der Aussage wurde diese bei 0 dB SNR nur mit 60 % der Urteile gewählt, während die übrigen Urteile für den Ruf abgegeben wurden. Die Verwechslungen sind weniger stark ausgeprägt als bei Satz Nr. 21 und betreffen vor allem die 0 dB SNR-Hörkondition.

Die Graphiken des Satzes Nr. 25 („Berlin ist eine große Stadt“) zeigen Schwierigkeiten der Probanden bei Darbietung der Kategorie Frage als auch der Kategorie Ruf. Für die Kategorie Frage liegen bei 0 dB SNR nur 60 % richtige Antworten vor, die übrigen 40 % wurden dem Ruf zugeordnet. Bei –6 dB SNR sind Verwechslungen mit der Aussage häufiger. Bei Darbietung der Kategorie Ruf kommt es ab –6 dB SNR zu einer Überkreuzung der Kurven von Ruf und Aussage. Die Graphik dieses Satzes verläuft insgesamt unterhalb der logistischen Funktion des Gesamttestsatzes.

Ferner zeigen die Sätze Nr. 6 („Der Fotoapparat funktioniert wieder“), Nr. 10 („In der Schule lernt man Mathematik“) und Nr. 23 („Der Koffer ist schwer“) Auffälligkeiten sowohl im Hinblick auf die Urteilsverteilung als auch auf die Gesamtanzahl der korrekten Antworten. Folgende Sätze sind in der vierten Spalte ebenfalls unterhalb der logistischen Funktion positioniert: Nr. 1, Nr. 12, Nr. 14, Nr. 32, Nr. 33 und Nr. 36. Die den genannten Satznummern entsprechenden Sätze sind in Anhang A aufgelistet. Bei diesen Sätzen sind die Abweichungen jedoch nur sehr gering ausgeprägt oder befinden sich vornehmlich im schwierigeren Bereich der Hörkonditionen.

Satz Nr. 15 („Das Konzert ist super“) ist exemplarisch für eine relativ treffsichere Zuordnung aller Kategorien durch die Probanden. Es liegt ein positiver Ausreißer vor.

#### 3.4.2 Erkennbarkeit der Sprecher

Bei den Sprechern sind im Gegensatz zu den Kategorien die Abweichungen wesentlich geringer ausgeprägt (s. Abbildungen C.7–C.12). Eine schlechtere Sprechererkennbarkeit bei einzelnen Sätzen in Relation zur Verstehbarkeit aller Sätze ist eher im Bereich der schwierigeren Hörkonditionen zu finden. Es sollen auch hier vier Beispiele genannt werden:

Bei Satz Nr. 5 („Mit dem Taschenrechner kannst du das“) sind bei Darbietung des männlichen Sprechers im mittleren Bereich der Hörkonditionen Verwechslungen mit der Frau erkennbar. Dies führt in der vierten Spalte zu einer Absenkung des Graphen im genannten Bereich.

Auch bei Satz Nr. 14 („Die Tagung war erfolgreich“) sind im mittleren bis schwierigen Bereich wenige richtige Antworten zu verzeichnen. Bedingt wird dies vor allem durch Verwechslungen der Frau mit dem Kind.

Satz Nr. 30 („Benzin ist wieder teurer geworden“) zeigt Ausreißer bei  $-6$  dB SNR. Hier wurde bei Darbietung der Frau relativ häufig das Kind genannt sowie der kindliche Sprecher insgesamt weniger erkannt.

Die Graphiken des Satzes Nr. 32 („Die Nachbarn sind laut“) ähneln denen des Satzes Nr. 30. Meist werden Frau und Kind verwechselt, wobei dies bei beiden Sätzen besonders die schwierigeren Hörkonditionen betrifft.

Als positive Ausreißer können bei den Sprechern Satz Nr. 18 („Du hast Besuch“) und Satz Nr. 19 („Das Buch ist spannend“) genannt werden. Bei Satz Nr. 18 wurde als einzigem der männliche Sprecher unter sämtlichen Hörbedingungen fehlerfrei erkannt. Ebenso weist die Frau auch im schwierigeren Hörbereich relativ viele korrekte Antworten auf. Auch Satz Nr. 19 fällt bei Darbietung des Mannes durch eine maximale Anzahl an richtigen Antworten bis hin zu  $-18$  dB SNR auf.

### **3.5 Empfindungen und Schwierigkeiten der Probanden während des Versuchs**

Einige Probanden stellten nach eingehender Instruktion die Frage, ob die verschiedenen Sprecher und Kategorien pro Liste gleich oft vorkämen. Dies beschreibt das schon vorab vermutete Problem (vgl. Kapitel 2.7). Sie erhielten die Anweisung, sich während des Versuchs ausschließlich auf das Gehörte zu konzentrieren. Trotzdem kann nicht ausgeschlossen werden, dass manche sich unbewusst an der von ihnen vermuteten Gleichverteilung orientierten.

Besonders die schwereren Hörkonditionen führten bei manchen Probanden vermehrt zu Zweifeln am eigenen Urteil, wenn sie öfter denselben Sprecher oder dieselbe Kategorie hintereinander zu hören glaubten. Laut eigener Aussage hätten sie aber versucht, sich im Zweifel auf den jeweils gehörten Satz zu konzentrieren.

Die Möglichkeit, sich vor Versuchsbeginn anhand eines Übungsdurchlaufs mit dem Test vertraut zu machen, wurde dankbar angenommen. Manche brauchten mehr Übungssätze und manche weniger, um sich für den Testablauf wirklich sicher zu fühlen. Dabei traten die meisten Schwierigkeiten bei der Zuordnung der Kategorien auf.

### *3.5 Empfindungen und Schwierigkeiten der Probanden während des Versuchs*

Auch nach Durchführung des Versuchs beschrieben fast alle Probanden den Kategorientest als schwieriger und anstrengender.

Die Probanden, die gleich zu Testbeginn mit dem lautesten Störgeräusch konfrontiert wurden, gaben das CCITT-Rauschen als unangenehmer an als jene, die erst gegen Ende des Tests das lauteste Rauschen hörten. Dabei gaben alle ein Gefühl der Frustration bei den lautereren Hörkonditionen an, da sie trotz intensivster Bemühungen oftmals nur wenig hörten.

Insgesamt waren alle Testpersonen sehr bemüht, möglichst viele richtige Antworten zu geben. Dabei fühlten sich manche unterfordert und hätten noch an weiteren Tests teilgenommen. Wieder andere hatten Mühe mit der Entscheidungsfindung und hätten im Zweifel lieber ein „Weiß nicht“ zur Antwort gegeben. Nur wenige fühlten sich eigenen Angaben zufolge nach einer Testdauer von eineinhalb bis zwei Stunden inklusive Tonaudiogramm am Rand ihrer Leistungsfähigkeit.

In Bezug auf die schallgedämpfte Hörkabine empfanden die meisten die ungewohnte Stille als angenehm. Nur wenige beschrieben den Raum als eher beängstigend und hatten zunächst ein beklommenes Gefühl.

Diese teilweise sehr unterschiedlichen Aussagen machen deutlich, welche unterschiedlichen individuellen Einflüssen die verschiedenen Versuchspersonen während des Tests unterlagen. Dies zu vermeiden, ist offensichtlich nicht oder nur sehr schwer möglich.

### 3 Ergebnisse

## 4 Diskussion

### 4.1 Betrachtung der Gesamtdaten

Ziel der vorliegenden Arbeit war die Durchführung eines Prosodietests mit normalhörenden Probanden in einem solchen Umfang, dass statistische Aussagen über die Eigenschaften des Tests getroffen werden konnten. Es sollten Schwachstellen aufgezeigt und wenn möglich beseitigt werden. Die verschiedenen lauten Störgeräusche sollten unterschiedliche Alltagssituationen und das damit verbundene prosodische Hörverstehen widerspiegeln. Zusätzlich könnten, mit den Resultaten der Normalhörenden als Vergleichswerten, hörgeschädigte Patienten sowie Hörhilfeträger anhand des Tests untersucht werden. Im nachfolgenden bestünde ebenfalls die Möglichkeit, verschiedene Signalverarbeitungsstrategien in Hörhilfen auf Aspekte der Prosodie zu untersuchen und weiterzuentwickeln.

Aus der Darstellung der Gesamtdaten in Abschnitt 3.1 ist ersichtlich, dass die Kurven aller Probanden im erwarteten Bereich von +10 dB bis –30 dB SNR liegen. Die Steilheit der Kurven – insbesondere für die Testart Kategorie – ergibt deutliche Wendepunkte, an denen das Hörverstehen der Probanden vom guten in den schlechten Hörbereich übergeht. Dies spricht ebenso wie die geringe Streuung der Einzelkurven für eine hohe Sensitivität des Tests. Im Gegensatz dazu kann bei kochleaimplantierten Patienten eine größere Varianz in der Lage der Kurven erwartet werden (Meister et al., 2007a). Umso wichtiger scheint eine geringe Variabilität der Ergebnisse bei Normalhörenden, da Hörgeschädigte oder Hörhilfeträger stets über ein unterschiedliches Hörvermögen verfügen und somit eine größere Varianz der Ergebnisse hervorrufen.

Die Graphik 3.1 zeigt, dass der Sprechergraph im Vergleich zur Kategorienkurve nach links verschoben ist. Der Wendepunkt des Graphen der Sprecher liegt mit –14,72 dB SNR links des Wendepunktes des Kategoriengraphen mit –11,65 dB SNR. Es liegt somit eine Verschiebung der Testart Sprecher in den Bereich der schwierigeren Hörkonditionen vor. Dies lässt erkennen, dass die Probanden bei der Zuordnung der Sprecher besser waren als bei der Zuordnung der Kategorien.

Bei 0 dB SNR lag die Anzahl der richtigen Antworten bei annähernd 100 %. Es

gab demzufolge keinerlei Schwierigkeiten bei der Differenzierung der unterschiedlichen Sprecher. Wurde die Hörkondition mit dem lautesten Störgeräusch dargeboten, waren die Antworten noch weit vom Zufallsniveau der drei Antwortmöglichkeiten von 33,3% entfernt. Es wurden noch zu fast 50% richtige Antworten gegeben. Die Probanden schienen also noch etwas zu hören, das sie in ihrer Entscheidung für die richtige Antwort beeinflusste.

Bei Darbietung der Kategorien erreichte die Kurve bei  $-24$  dB SNR annähernd die 33,3%-Marke. Die Versuchsteilnehmer gaben im gesamten Kurvenverlauf schlechtere Antworten. Bei 0 dB SNR lag die Anzahl der Treffer nur bei ca. 95%. Die Ursache hierfür könnte eher in Interpretationsfehlern als in Fehlern der Prosodieperzeption zu suchen sein. Auch [Meister et al. \(2007a\)](#) erhielt bei seinen Tests mit natürlich-sprachlichen Aufnahmen in ruhiger Umgebung das Resultat, dass alle Urteile für die Sprecher korrekt ausfielen, während hingegen bei den Satzmodi in geringem Maß Verwechslungen auftraten. Es stellt sich die Frage, ob die Interpretationsfehler bestimmten Sätzen zugeordnet werden können. Dies wurde bereits in Abschnitt 3.4 eingehend untersucht und soll im noch folgenden Abschnitt 4.4 diskutiert werden.

Bei der Erkennung des Sprechergeschlechts spielt die Intonation als linguistischer Aspekt der Prosodie eine bedeutende Rolle. Die Intonation ist eng mit der *absoluten Lage* des akustischen Parameters Grundfrequenz verknüpft. Die drei Sprecher besitzen verschiedene Grundfrequenzen (vgl. Abschnitt 4.3).

Die Kategorienerkennung basiert hingegen eher auf dem *Verlauf* der Grundfrequenzen. So ist bei der Frage ein Anstieg auf den letzten Silben zu verzeichnen, während die Grundfrequenz bei einer Aussage gegen Ende des Satzes eher abfällt. Außerdem spielt der Grundfrequenzverlauf auch bei betonten Silben eine Rolle, da hier die Stimme typischerweise angehoben wird. Als akustische Charakteristika einer betonten Silbe sind zudem noch die Amplitude und Veränderungen der Dauer zu vermerken. Betonungen sind besonders beim Ruf als auch bei der Frage, die beide als emotionsgeladen zu betrachten sind, bedeutsam.

Es scheint den Probanden leichter zu fallen, zwischen den verschiedenen Grundfrequenzen der Sprecher zu differenzieren als aus Änderungen im Verlauf einer Grundfrequenz auf die zugehörige Kategorie zu schließen.

Der steilere Verlauf des Kategoriengraphen spiegelt die schnellere Abnahme der Fähigkeit zur Kategorienerkennung durch die Probanden wider. Offensichtlich konnten sie ab einer bestimmten Lautheit des Rauschens die Grundfrequenzverläufe nicht mehr identifizieren. Gewisse hörbare Wortspitzen waren nicht mehr zuzuordnen. Hingegen ist es wahrscheinlich, dass diese Wortspitzen beim Sprechertest

mehr Informationen lieferten. Es scheint einfacher zu sein, hörbare Bruchstücke eines Satzes der Grundfrequenz per se und somit dem Sprechergeschlecht zuzuordnen als den Verlauf der Grundfrequenz oder Wortakzente zu beurteilen. So konnte beispielsweise bei einer hohen Tonlage der Mann ausgeschlossen werden und das Zufallsniveau der Antwortmöglichkeiten auf 50 % abgesenkt werden.

Die Einzelkurven für die Testart Kategorie in Abbildung 3.3 auf Seite 42 zeigen eine geringere Streuung als die des Sprechertests. Dies ist u. a. auf den flacheren Verlauf des Sprechertests zurückzuführen.

## 4.2 Betrachtung der Urteilsverteilung

### 4.2.1 Urteilserkennung

Aus dem Blasendiagramm 3.4 (s. Seite 44) ist ersichtlich, wie gut oder schlecht die Versuchsteilnehmer die dargebotenen Sprecher und Kategorien bei unterschiedlichen Hörbedingungen erkannten. Zu erwarten wäre eine ähnlich gute Erkennung aller Sprecher bzw. Kategorien, deren Trefferquoten von einer weitgehend fehlerfreien Zuordnung in den leichten Hörkonditionen im SNR-abhängigen Verlauf kontinuierlich und gleichmäßig abnehmen. Entgegen dieser Erwartungen fiel die Abnahme der Urteilserkennung sowohl von den Sprechern als auch den Kategorien ungleichmäßig aus. Dies ist beim Sprechertest besonders ausgeprägt und soll an diesem zuerst diskutiert werden.

Bei Darbietung der leichtesten Hörkondition wurden alle Sprecher nahezu fehlerfrei erkannt. Die Probanden waren demzufolge in der Lage, bei nur leisem Hintergrundgeräusch mühelos zwischen den dargebotenen Sprechern zu differenzieren. Die ungleichmäßige Abnahme der Treffer im SNR-abhängigen Verlauf erfolgt vor allem zu Gunsten des Mannes und zu Lasten des Kindes, das bei  $-24$  dB SNR nur etwa ein Viertel der Treffer des Mannes aufweist. Bei diesen Beobachtungen muss berücksichtigt werden, dass die Urteilserkennung im SNR-abhängigen Verlauf durch die Annäherung an die für jeden Sprecher und jede Kategorie spezifische Ratewahrscheinlichkeit beeinflusst wird (s. Abschnitt 4.3). Des weiteren könnten Verwechslungen sowie „umgekehrte Treffer“ mögliche Einflussfaktoren sein. Über Ursachen von Verwechslungen soll weiter unten spekuliert werden (s. Abschnitt 4.2.2).

Für die Urteilserkennung der Kategorien gelten die gleichen Überlegungen. Hierbei ist jedoch auffallend, dass die Trefferquoten bei Darbietung der leichtesten Hörkondition etwas unterhalb derjenigen des Sprechertests platziert sind. Dies ist vor

allem auf Verwechslungen von Aussage und Ruf zurückzuführen und soll weiter unten erläutert werden (s. Abschnitt 4.2.2). Im weiteren SNR-abhängigen Verlauf nimmt die Anzahl der korrekt erkannten Kategorien stärker aber auch gleichmäßiger ab als die des Sprechertests. Die stärkere Abnahme der Trefferanzahl spricht für eine insgesamt schlechtere Zuordnung der Kategorien. Mögliche Ursachen hierfür wurden bereits in Abschnitt 4.1 erläutert. Die gleichmäßigere Abnahme der erkannten Urteile, die dennoch eine deutliche Tendenz zu Gunsten der Aussage und zu Lasten der Frage aufweist, kann mit einer Annäherung an die Ratewahrscheinlichkeiten besonders bei den lautesten Hörkonditionen (s. Abschnitt 4.3) und Verwechslungen (s. Abschnitt 4.2.2) erklärt werden.

### 4.2.2 Verwechslungen

Im Sprechertest wurde der Mann im leichten Hörbereich gar nicht und unter mittelschweren Hörbedingungen eher mit der Frau verwechselt. Dies könnte mit den unterschiedlichen Grundfrequenzen der Sprechstimmen erklärt werden. So liegt die Grundfrequenz des Mannes mit 96,45 Hz näher bei der der Frau mit 190,7 Hz als bei der des Kindes (226,3 Hz). Noch hörbare Wortspitzen könnten so irrtümlich der Frau zugeordnet worden sein. Ab  $-18$  dB SNR stiegen jedoch Verwechslungen mit dem Kind ebenfalls an. Mögliche Ursachen könnten in dem Umstand zu suchen sein, dass das Störgeräusch bei  $-18$  dB SNR mehr Wortspitzen überlagert und die Probanden so vermehrt auf Raten angewiesen waren. Manche könnten auch bei Darbietung der leichteren Hörkonditionen die jeweils gleiche Anzahl der verschiedenen Sprecher pro Liste bemerkt haben und ihre Urteile gleichmäßiger verteilt haben. Bei Darbietung des weiblichen und des kindlichen Sprachstimulus traten schon bei relativ leisem Rauschen „umgekehrte Treffer“ auf, was vermutlich auf die ähnlichen Grundfrequenzen zurückzuführen ist. In der leichtesten Hörkondition wurden bereits fünf Urteile dem Kind anstelle der Frau zugeordnet. Im SNR-abhängigen Verlauf nahmen Verwechslungen von Frau und Kind mit dem Mann als Annäherung an die Ratewahrscheinlichkeiten bedeutend zu (s. Abschnitt 4.3).

Bei der Testart Kategorie traten vorwiegend „umgekehrte Treffer“ zwischen Aussage und Ausruf auf. Da diese Verwechslungen schon bei der 0 dB SNR-Hörbedingung existierten, sind Interpretationsfehler der Testpersonen zumindest unter den leichtesten Hörbedingungen wahrscheinlich. In Abschnitt 3.4 wurden einige Sätze beschrieben, bei denen es öfter zu Verwechslungen kam und die somit unglaubwürdiger und für den Test weniger geeignet scheinen. Es besteht daher die Möglichkeit, dass die



Verwechslungen auf diesen Sätzen basieren. Die Ursache für diese Interpretationsfehler könnte in den zugrunde liegenden akustischen Parametern zu suchen sein. Wie bereits erwähnt, zeichnet sich die Frage als einzige Kategorie durch einen Anstieg der Grundfrequenz auf den letzten Silben eines Satzes aus. Bei der Aussage und dem Ruf hingegen sinkt die Grundfrequenz gegen Ende des Sprachstimulus ab. Denkbar wären demzufolge Verwechslungen der beiden Kategorien aufgrund der ähnlichen Grundfrequenzverläufe. Man könnte nun einwenden, dass sich andererseits Frage und Ruf durch Akzente und Emotion von der trockenen und emotionslosen Aussage unterscheiden. Dies wirft die Frage auf, warum nicht Frage und Ruf eine Vielzahl von Verwechslungen aufweisen. Bei genauer Betrachtung der 0 dB SNR-Graphik des Blasendiagramms 3.4 ist auffallend, dass bei Darbietung der Frage tatsächlich mehr Verwechslungen mit dem Ruf erfolgten. Da jedoch meist Aussage und Ruf zu Gunsten der Aussage verwechselt wurden, besteht die Möglichkeit einer besseren Hörbarkeit und Prägnanz des Grundfrequenzverlaufs als der Intensität der Akzente.

In der Literatur finden sich meist Experimente, in denen ausschließlich Aussage und Frage als Ausdruck des Grundfrequenzverlaufs als Satzmodi untersucht werden. [Meister et al. \(2007b\)](#) testete z. B. in den beiden Untertests Satzmodi (Aussage vs. Frage) und Satzakkente Normalhörende und CI-Träger. Sowohl die normalhörenden Probanden als auch die CI-Patienten schnitten im Untertest Satzmodi besser ab, wobei die interindividuelle Streuung bei den CI-Nutzern größer war. Man könnte nun spekulieren und die oben aufgestellte Vermutung, dass Grundfrequenzverläufe in der Perzeption einfacher sind als Akzente, als wahrscheinlich betrachten. Ob diese Überlegungen jedoch ebenfalls auf einen Vergleich zwischen Aussage und Ruf in der vorliegenden Arbeit übertragbar sind, ist auf Grund methodischer Unterschiede fraglich.

### 4.2.3 Urteilshäufigkeit

Zu erwarten wäre eine durchgängig gleiche Urteilshäufigkeit mit jeweils ca. 360 Urteilen pro Sprecher oder Kategorie gewesen. Bedingt wäre dies durch richtiges Antworten in den leichteren und durch totales Raten in den schwereren Hörkonditionen. Ab  $-12$  dB SNR scheint die Anzahl der erkannten Sprecher oder Kategorien stärker abzunehmen (s. Abschnitt 3.1 und 3.4). Die Antworten sind jedoch nicht gleichmäßig auf die dargebotenen Sprecher oder Kategorien verteilt, sondern nähern sich im SNR-abhängigen Verlauf bestimmten Werten an. Diese Werte entsprechen für die

schwerste Hörkondition den weiter unten beschriebenen Ratewahrscheinlichkeiten. Die Urteilshäufigkeit im SNR-abhängigen Verlauf kann demzufolge als eine Annäherung der Antworten von einer maximalen Trefferquote bis hin zu totalem Raten gesehen werden.

### 4.3 Betrachtung der Ratewahrscheinlichkeiten

Bei Betrachtung der Funktionen der Urteilsverteilung in Abschnitt 3.3 ist auffallend, dass sich diese im SNR-abhängigen Verlauf über die  $-24$  dB-Marke hinaus unterschiedlichen Werten annähern. Diese Werte stellen die Leistungen der Probanden bei totalem Raten dar. Zu erwarten wäre eine Gleichverteilung der Urteile gewesen, die in Werten von etwa 33,3 % pro Urteil resultierte. Entgegen dieser Erwartungen wurde bei der Testart Kategorie die Aussage mit 45,7 % weit häufiger als die Frage mit 23,1 % gewählt. Beim Sprechertest treten die Unterschiede in der Urteilsverteilung noch deutlicher hervor. Die Hälfte der Urteile wurden dem Mann zugesprochen, die Ratewahrscheinlichkeit für den Sprecher Kind beträgt nur 17,8 %. Die Leistungen der Probanden bei totalem Raten sind in Tabelle 3.1 auf Seite 47 abgebildet. Hinsichtlich der unterschiedlichen Ratewahrscheinlichkeiten stellt sich die Frage, welche Einflussfaktoren dem zugrunde liegen könnten.

Ausgehend von der Annahme, dass die Probanden bei Darbietung der schwierigsten Hörkondition den Sprachstimulus nicht mehr wahrnahmen und folglich jedes Hörverstehen auszuschließen wäre, besteht die Möglichkeit, dass die ungleiche Verteilung der Ratewahrscheinlichkeiten einfach Folge des Rateverhaltens der Versuchsteilnehmer ist. Einflussfaktor wäre somit nur das Raten per se. Eine mögliche Ursache des Rateverhaltens könnte in der Geläufigkeit der unterschiedlichen Kategorien und Sprecher zu suchen sein. Die Probanden dachten bei der Entscheidungsfindung für ein Urteil eventuell einfach zuerst an die Aussage bzw. den Mann und gaben dementsprechend ihr Urteil ab. Unterstützt wird diese Hypothese durch die Tatsache, dass vor Versuchsbeginn Schautafeln mit allen Antwortmöglichkeiten ausgehändigt wurden. Dadurch sollte eine ständige Vergegenwärtigung der drei möglichen Urteile pro Testart erfolgen und verhindert werden, dass eventuell ein Sprecher oder eine Kategorie im Testverlauf in Vergessenheit gerät und bei der Urteilsabgabe außer Acht gelassen wird. Die Reihenfolge der Urteile auf den Schautafeln von links nach rechts war folgende: Aussage – Frage – Ausruf für den Kategorientest bzw. Mann – Frau – Kind für die Sprecher. Entsprechend der Lesegewohnheiten der

Probanden von links nach rechts wären dementsprechende Einflüsse in der Urteilsabgabe möglich.

Als weitere mögliche Ursache für das Rateverhalten per se sind Einflüsse durch die gewohnte Häufigkeit der unterschiedlichen Antwortmöglichkeiten im Alltag denkbar. Diese Überlegung betrifft v. a. die Testart Kategorie. So könnten im alltäglichen Leben der Versuchsteilnehmer Aussagen häufiger vorkommen als Rufe oder Fragen, so dass die Probanden durch die mögliche häufigere Verwendung von Aussagen im Sprachgebrauch diese eher gewohnt sind. Diese „gewohnte Häufigkeit“ könnte womöglich zu einer bevorzugten Assoziation des Gehörten mit der Aussage geführt haben.

Die genannten Einflussfaktoren wurden unter einem Gesichtspunkt diskutiert, der jegliche Einflüsse durch CCITT-Rauschen sowie Sprachstimuli ausschließt. Dieser Fall läge aber nur dann vor, wenn weder eine Darbietung des CCITT-Rauschens noch der Sprachstimuli erfolgt wäre. Es kann daher nicht ausgeschlossen werden, dass das Rauschen selber einen Einflussfaktor darstellt.

Um mögliche Einflüsse zunächst nur durch die Sprachstimuli zu betrachten, muss auf die unterschiedlichen zugrunde liegenden Grundfrequenzen eingegangen werden. Die Grundfrequenz des männlichen Sprechers liegt bei 96,45 Hz, die der Frau bedeutend höher bei 190,7 Hz. Das Kind spricht mit der höchsten Grundfrequenz von 226,3 Hz. Folglich sprechen Frau und Kind mit ähnlicheren Grundfrequenzen. [Lehiste und Meltzer \(1973\)](#) untersuchten die Erkennung von Sprechern mit Hilfe natürlich gesprochener und synthetisierter Vokale. Sie beobachteten, dass männliche Sprecher leichter als Frauen und Kinder erkannt wurden, die öfter Verwechslungen ausgesetzt waren. Dabei waren die Verwechslungen bei den synthetisierten Stimuli stärker ausgeprägt. In einem dritten Subtest verwendeten [Lehiste und Meltzer \(1973\)](#) ebenfalls synthetisierte Vokale, bei deren Erzeugung jedoch die unterschiedlichen charakteristischen Grundfrequenzen von Mann, Frau und Kind mit männlichen, weiblichen sowie kindlichen Formanten kombiniert wurden. Die Ergebnisse zeigen, dass Formantstrukturen wichtigere Hinweise bei der Sprechererkennung zu liefern scheinen als die Grundfrequenz. So wurden z. B. mit männlichen Formanten und weiblicher Grundfrequenz erzeugte Vokale zu ca. 81 % dem männlichen Sprecher zugeordnet, während Vokale mit weiblichen Formanten und männlicher Grundfrequenz dem Mann nur zu ca. 19 % zugeordnet wurden. [Darwin et al. \(2003\)](#) beschreiben bei der Identifikation des Sprechergeschlechts ebenfalls einen Einfluss der unterschiedlichen Formantfrequenzen neben den Grundfrequenzen. Sie stützen sich auf Studien von [Peterson und Barney \(1952\)](#), denen zufolge die Formantfrequenzen von

## 4 Diskussion

Frauen um etwa 16 % höher liegen als die der Männer. [Fu et al. \(2005\)](#) nennen Unterschiede von ca. 15 % zwischen weiblichen und männlichen Sprechern. Die Varianz der Formantfrequenzen erklärt sich mit unterschiedlichen anatomischen Formen des Vokaltrakts. In keiner der letztgenannten Studien ist jedoch ein kindlicher Sprecher involviert. In den genannten Studien werden neben Änderungen der Grundfrequenz die verschiedenen Formantfrequenzen als wichtige Kriterien bei der Sprechererkennung betrachtet, so dass diese ebenfalls Beachtung finden sollten.

Als Neben aspekt sollen Beobachtungen von [Lehiste und Meltzer \(1973\)](#) erwähnt werden, denen zufolge von Kindern gesprochene Vokale schwieriger zu verstehen sind. Die Schwierigkeit könnte auf den ersten Blick mit einer undeutlichen „vernuschelten“ Aussprache und somit unklaren prosodischen Gestaltung der Sätze während der Aufsprache erklärt werden. Vergleiche mit Untersuchungen von [Lehiste und Meltzer \(1973\)](#) und [Peterson und Barney \(1952\)](#) zeigen jedoch, dass dies nicht zutreffend ist. Die genannten Autoren ermittelten ebenfalls die niedrigsten Trefferquoten für kindliche Vokale. [Lehiste und Meltzer \(1973\)](#) vermuten die Ursache hierfür in Beobachtungen, dass die Formanten eines Kindes aufgrund der hohen Grundfrequenz nicht deutlich artikuliert werden und die Vokale daher schlechter erkennbar sind. Diese Beobachtungen betreffen jedoch lediglich die Spracherkennung und sollten bei Überlegungen in Bezug auf die Sprecheridentifikation keinen Einflussfaktor darstellen.

Neben Eigenschaften der Sprachstimuli als solchen können aber auch Eigenschaften des Störgeräuschs eine Rolle gespielt haben. Charakteristische Merkmale des sprachsimulierenden Rauschens wurden in Abschnitt [2.3.2](#) bereits erläutert. Der Frequenzbereich erstreckt sich von etwa 90 Hz bis 900 Hz. Demzufolge liegt der Mann eher im unteren Frequenzbereich des Rauschens, während das Kind im mittleren Bereich zu finden ist. Entsprechend der unterschiedlichen Lage der Grundfrequenzen der Sprecher im Spektrum des Breitbandrauschens besteht die Möglichkeit, dass die Versuchsteilnehmer die Identifikation der verschiedenen Sprecher unterschiedlich schwierig fanden. Der Mann im unteren Bereich des Frequenzspektrums wurde am häufigsten erkannt. Bedingt durch die grenzwertige Lage „fiel dieser etwas aus dem Rauschen heraus“ und war aufgrund größeren SNRs so möglicherweise günstiger zu beurteilen und leichter zuzuordnen. Die Identifikation der Frau gelang deutlich weniger, während das Kind im Zentrum des Spektrums mit den wenigsten Treffern zugeordnet werden konnte.

Die unterschiedlichen Ratewahrscheinlichkeiten für die Testart Sprecher können so erklärt werden. Dies ist jedoch für den Kategorientest, dessen Satzmaterial aus-

schließlich vom Mann gesprochen wird, nicht zutreffend. Hier muss nach anderen Erklärungen gesucht werden. Während bei der Identifikation der Sprecher die Grundfrequenz per se dominierendes Kriterium ist, spielt beim Kategorientest neben den akustischen Parametern Intensität und Dauer der Grundfrequenzverlauf eine wichtige Rolle. Dieser wird bei der Aussage am Satzende typischerweise abgesenkt und bei der Frage angehoben. Die Ratewahrscheinlichkeiten deuten darauf hin, dass ein Anstieg der Tonhöhe auf den letzten Silben eines Satzes schwieriger zu hören ist, zumal höhere Frequenzen stets vermehrt vom Störgeräusch überlagert werden. Diese Annahme deckt sich mit Angaben aus der Literatur. [Meister et al. \(2007a\)](#) modifizierte natürlichsprachliche Aufnahmen mittels einer Sprachsignalverarbeitungssoftware, wodurch neun Stufen zwischen den jeweiligen Extrema (z. B. Aussage und Frage bei den Satzmodi) erzeugt wurden. Er fand heraus, dass ein leichtes Anheben der Grundfrequenz auf der letzten Silbe eher noch als Aussage wahrgenommen wurde und erst ein deutlicher Anstieg einer Frage zugeordnet wurde. Eine Begründung hierfür ist nach [Meister et al. \(2007a\)](#) in dem Umstand zu suchen, dass die Grundfrequenz grundsätzlich auch ein wichtiger Indikator für Phrasengrenzen ist. Experimente von [Darwin et al. \(2003\)](#) zeigten eine deutliche Wahrnehmung von Unterschieden in der Grundfrequenz erst ab zwei Halbtönen. Das Hauptentscheidungsmerkmal der Kategorie Frage wird demzufolge in seiner Identifikation erschwert. Fraglich ist jedoch, inwieweit Tonhöhenänderungen desselben Sprechers durch ihre Lage im Spektrum des Breitbandrauschens überlagert werden und inwiefern dies für die Entscheidung zwischen Aussage und Frage relevant ist.

Denkbar wäre auch eine Mehrzahl der abgegebenen Urteile für die Aussage aufgrund eines Verlustes des Hörverstehens der Probanden. Waren weder Betonungen, wie dies bei den Kategorien Ruf und Frage charakteristisch ist, noch ein Anheben der Tonhöhe am Satzende wie bei der Frage wahrnehmbar, wurde das Urteil Aussage gefällt. Dieses Verhalten der Versuchsteilnehmer beruhte auf zwei Ursachen. Zum einen handelte es sich tatsächlich um die Aussage und die Fällung des Urteils war korrekt. Andererseits ist es denkbar, dass trotz Darbietung des Rufs oder der Frage die entsprechenden charakteristischen Merkmale nicht gehört wurden. Die Möglichkeit einer eventuellen schlechteren Perzeption der Akzente (wie bei einer Frage oder einem Ruf) als der Grundfrequenzverläufe wurde bereits in Abschnitt [4.2.2](#) diskutiert. Auch über eine mögliche schlechtere Wahrnehmung des Anstiegs der Grundfrequenz auf der letzten Silbe wurde weiter oben spekuliert. Demzufolge könnte fälschlicherweise von einer Aussage ausgegangen worden sein. Diese Hypothese wird durch Anmerkungen einiger Probanden unterstützt. Ihnen zufolge hatten

sie bei totalem Verlust des Hörverstehens immer die Aussage gewählt.

Als weiterer denkbarer Einflussfaktor sollte die Gleichverteilung der Kategorien bzw. Sprecher innerhalb der Listen berücksichtigt werden. Jede Antwortmöglichkeit war in jeder Liste gleich oft vertreten, was den Versuchsteilnehmern nicht mitgeteilt worden war. Dies könnte theoretisch bei den leichteren Hörkonditionen bemerkt worden sein und Einfluss auf das Antwortverhalten bei schwierigeren Hörbedingungen genommen haben. Im Widerspruch zu dieser Hypothese steht jedoch die Beobachtung, dass die Antworten der Probanden eher durch eine ungleichmäßige Verteilung geprägt waren. Unklar bleibt jedoch, ob bei totalem Fehlen dieses Einflussfaktors eventuell eine noch größere Ungleichmäßigkeit im Antwortverhalten zu beobachten gewesen wäre. Folglich sollte die genannte Hypothese nicht völlig außer acht gelassen werden.

Es stellt sich die Frage, ob manche der genannten Einflüsse eine stärkere Rolle bei der Urteilsabgabe spielten als andere. Unklar bleibt auch die Abhängigkeit der Faktoren voneinander. Dies ist jedoch nicht oder nur sehr schwierig zu ermitteln, da vermutlich probandenspezifische Unterschiede beruhend auf individuellen Hörgewohnheiten, prosodischem Sprachverständnis sowie Intellekt existierten (vgl. Abschnitt 4.5). Des Weiteren sind andere der Verfasserin dieser Arbeit unbekanntere Einflüsse denkbar.

### 4.4 Betrachtung der Verstehbarkeit der Sätze

In Abschnitt 3.4 wurden einige Ausreißersätze genannt, die für den Test weniger geeignet scheinen. Es stellt sich nun die Frage, worin die Ursache hierfür liegen könnte und worauf bei der Konstruktion neuer verbesserter Sätze geachtet werden sollte.

Bei der Testart Kategorie sind die Ausreißer ausgeprägter und treten eher bei den leichteren Hörkonditionen auf als beim Sprechertest, der bei Darbietung der Sätze unter leichten Hörbedingungen kaum Abweichungen aufweist. Dies entspricht unseren Erwartungen, gemäß denen der Sprechertest weitgehend unabhängig von der Konstruktion der einzelnen Sätze sein und ausschließlich auf der Varianz der Sprechstimmen beruhen sollte. Demzufolge liegt der Diskussionsschwerpunkt der möglichen Ursachen auf der Testart Kategorie. Bedingt durch die Grundstruktur der Sätze besteht die Möglichkeit, dass manche Sätze in einer bestimmten Kategorie seltsam und unglaubwürdig klingen. So kann beispielsweise ein Satz in Form der Aussage

normal, als Frage jedoch seltsam klingen, so wie man ihn im normalen Sprachgebrauch nicht verwenden würde. Dies kann zu Schwierigkeiten im Hörverstehen der Probanden und somit zu Fehlinterpretationen führen. Es ist relativ unwahrscheinlich, dass bestimmte Sätze bei der Sprechererkennung schwieriger sind als andere. Zudem weist der Sprechertest eine größere Varianz in der Punktverteilung auf, so dass der zugehörige Graph der Gesamtheit aller Sätze beim Vergleich mit den Kurven der Einzelsätze eventuell zu Fehlinterpretationen führen könnte.

##### 4.4.1 Erkennbarkeit der Kategorien

Zunächst sollen mögliche Ursachen für Ausreißer der Testart Kategorie erläutert werden. Die vier Sätze „Schön haben wir es hier“ (Nr. 3), „Der Maler ist wirklich begabt“ (Nr. 21), „Klettern kann gefährlich sein“ (Nr. 22) und „Berlin ist eine große Stadt“ (Nr. 25) fallen am stärksten aus dem Raster (s. Abschnitt 3.4). Alle der genannten Sätze enthalten mehr oder weniger stark emotionsgefärbte Wörter, bei denen es sich meist um Adjektive oder Adverbien handelt. So beschreibt das „schön“ in Satz Nr. 3 die subjektive Ansicht des Sprechers, es impliziert eine Wertung. Um das Wort schön zu gebrauchen, muss er sich in einer bestimmten Gemütsverfassung befinden. Auf diese Weise erhält der Satz allein durch Verwendung dieses einen Wortes eine emotionale Färbung und könnte so bei einigen Probanden eine Assoziation mit dem Ruf bewirkt haben, dessen Hauptkennungsmerkmal „gesprochene“ Emotion und Akzentuierung ist. Tatsächlich traten bei Satz Nr. 3 meistens Verwechslungen von Aussage und Ruf auf. Bei akustischer Darbietung des Satzes ist auffällig, dass das Adverb „schön“ bei allen drei Kategorien stets Akzentuierungen aufweist. Dies ist durch den Satzbau bedingt und könnte ebenfalls dazu beigetragen haben, die Differenzierung zwischen stärkeren Betonungen wie beim Ruf und schwachen natürlichen Hervorhebungen wie bei der Aussage zu erschweren.

Auch in den Sätzen Nr. 21 („Der Maler ist wirklich begabt“) und Nr. 22 („Klettern kann gefährlich sein“) schwingt durch die Wortwahl bedingt Emotion mit. So drückt das Wort „wirklich“ eine Überzeugung des Sprechers hinsichtlich der Begabung des Malers aus, während mit Hilfe des Wortes „gefährlich“ in Satz Nr. 22 eine Warnung ausgesprochen wird. Hierdurch wird die Zuordnung zu einer emotionslosen und eher objektiven Aussage erschwert. Dies zeigt sich ebenfalls in der Beobachtung, dass die beiden letztgenannten Sätze bei Darbietung der Kategorie Aussage in der entsprechenden Graphik eine starke Tendenz zum Urteil Ruf aufweisen.

Diese Hypothese wird durch die Beobachtung unterstützt, dass die größten Ab-



weichungen der Ausreißersätze im Bereich der leichten Hörkonditionen vorliegen, in denen einzelnen Wörter noch verständlich waren.

Eine Ausnahme bildet Satz Nr. 15 („Das Konzert ist super“), der in jeder Kategorie gut erkannt wurde und damit einen Positivausreißer darstellt. Aus der graphischen Darstellung im Anhang C.3 ist jedoch ersichtlich, dass die Zuordnung des Satzes am besten für die Kategorie Frage, aber auch für den Ruf sehr gut gelang. Dies ist für den Ruf mit dem emotionsgeladenen Wort „super“ und dessen Betonung im gesprochenen Satz erklärbar. Auch bei der Frage, die Emotionen beinhalten kann, ist dies zutreffend. Zudem ist dieser Satz in den Kategorien Frage und Ausruf allgemein geläufig und wird oft verwendet. Er klingt weder seltsam noch unglaubwürdig. Bei den Sätzen Nr. 10 („In der Schule lernt man Mathematik“) und Nr. 25 („Berlin ist eine große Stadt“) ist dies nicht der Fall. Beide sind in Form der Frage und des Rufs seltsam formuliert, so wie sie im normalen Sprachgebrauch nicht verwendet werden. Die Sätze Nr. 10 sowie Nr. 25 wurden demzufolge weniger aufgrund emotionstragender Wörter schlechter erkannt, sondern eher bedingt durch ihre Satzstruktur.

Eine plausible Erklärung für die ebenfalls gute Zuordnung der Aussage von Satz Nr. 15 kann nicht gegeben werden. Mögliche Ursachen könnten in der Aussprache der Kategorien, die durch besonders prägnante Akzentuierungen eine Differenzierung zwischen Aussage, Frage und Ruf erleichterte oder dem simplen Satzbau zu suchen sein. Leichte Akzentuierungen bestimmter Wörter oder Phrasen sind auch bei der Aussage vorhanden und notwendig, um einen gesprochenen Satz natürlich klingen zu lassen. In der Aussage des Satzes Nr. 15 liegt der Schwerpunkt auf dem Subjekt „Konzert“ und nicht wie beim Ruf auf „super“. So könnten die Versuchsteilnehmer durch die Nichtbetonung des Wortes „super“ auf eine Aussage geschlossen haben.

### 4.4.2 Erkennbarkeit der Sprecher

Da in geringerem Ausmaß ebenfalls Abweichungen beim Sprechertest auftreten, sollen mögliche Ursachen hierfür an dieser Stelle diskutiert werden. Negative Ausreißer sind vorwiegend im Bereich der schweren und mittleren Hörkonditionen positioniert, während es bei 0 dB SNR kaum Verwechslungen gab.

Da bei den leichteren Hörbedingungen nur wenige Fehlantworten vorliegen, scheint es eher unwahrscheinlich, dass es sich um durch den Satzbau und -inhalt bedingte Interpretationsfehler handelt. Es könnten eher Konzentrationsschwächen oder mangelhafte Motivation der Probanden als Ursachen in Betracht gezogen werden. An-



dererseits kann nicht ausgeschlossen werden, dass manche Sätze aufgrund ihres Satzinhalts eventuell für einen oder mehrere Sprecher untypisch sind. So werden im deutschen Sprachgebrauch bestimmte Wörter wie z. B. „Haushalt“ oder „Kochbuch“ öfter von weiblichen als von den anderen beiden Sprechern verwendet. Begriffe wie „Auto“ oder „Fußball“ werden meist eher mit einem Mann assoziiert und wieder andere sind typisch für ein Kind. Es besteht die Möglichkeit, dass Probanden beim Hören einzelner Wortspitzen solche „Schlagworte“ unbewusst mit einer ihrer Meinungen nach für sie typischen Sprecher in Verbindung brachten und dementsprechend urteilten. Eine Probandin hörte beispielsweise bei Darbietung des Satzes „Benzin ist wieder teurer geworden“ nur das Wort „Benzin“ und stellte nach Durchführung des Tests die Frage, ob der genannte Satz auch von einem Kind gesprochen worden sein könnte. Als Urteil hatte sie einen der anderen beiden Sprecher angegeben, da sie das Kind für unwahrscheinlich hielt. Des Weiteren ist das jeweils individuelle soziale Umfeld eines jeden Probanden und demzufolge mögliche unterschiedliche Assoziationen mit bestimmten Sprechern zu berücksichtigen.

Neben oben genanntem Satz „Benzin ist wieder teurer geworden“ könnten die Sätze „Die Vermieter sind freundlich“ sowie „Die Tagung war erfolgreich“ als untypisch für einen kindlichen Sprecher angesehen werden. Diese Überlegungen lassen sich im Test nur teilweise verifizieren. Eventuelle Übereinstimmungen könnten ebenfalls Folge anderer unbekannter Ursachen sein und müssen daher kritisch gesehen werden. Dennoch sollte bei der Entwicklung neuer Sätze dieses Problem berücksichtigt werden, um mögliche Einflüsse auszuschalten.

Im Bereich der schwierigeren Hörkonditionen konnte der Satzinhalt nicht mehr gehört werden. Somit kann die Konstruktion des Satzes per se als Ursache für negative Ausreißer nicht in Betracht gezogen werden. Anders verhält es sich mit den Frequenzbereichen bestimmter Vokale und Zischlaute. Ruhe (1999) beispielsweise schreibt: „Während sich die Grundtonhöhe beim Sprechen nur wenig ändert, bestehen in Abhängigkeit vom jeweils gesprochenen Phonem ganz erhebliche Unterschiede in den Klangfarben.“ Der Hauptsprachbereich der Vokale und Konsonanten deckt einen Bereich zwischen 500 und 2000 Hz ab, während die Grundfrequenzen der verschiedenen Sprecher bei 100 bis 250 Hz liegen. Der Mann mit einer Grundfrequenz von 96,45 Hz wurde am häufigsten erkannt. Es besteht die Möglichkeit, dass Sätze mit dunkel klingendem /u/ aufgrund des tieferen Frequenzanteils eher dem männlichen Sprecher zugeordnet wurden. Ruhe (1999) beschreibt die Hauptenergieanteile von /u/ und /i/ folgendermaßen: „[...] /u/ hat z. B. die Hauptenergieanteile zwischen 200 Hz und 600 Hz (1. Formant), während beim wesentlich heller klingenden Vokal /i/

hochfrequente Anteile zwischen 2000-5000 Hz hinzutreten (2. Formant).“ Zischlaute weisen ebenfalls hochfrequente Anteile auf. Die Hauptenergie des Phonems /sch/ liegt zwischen 1600 und 8000 Hz, die des noch schärfer klingenden /tz/ bei 5000-15000 Hz.

Auf diese Weise könnte man Abweichungen in den mittelschweren Hörkonditionen des Satzes Nr. 5 („Mit dem Taschenrechner kannst du das“) erklären. Es treten vermehrt scharf gesprochene Laute auf, die aufgrund ihrer hochfrequenten Anteile zu Verwechslungen mit einer Sprechstimme mit höherer Grundfrequenz geführt haben könnten. Tatsächlich wurde bei Darbietung des Mannes im genannten Hörbereich relativ häufig die Frau gewählt.

Die beiden Sätze „Du hast Besuch“ (Nr. 18) sowie „Das Buch ist spannend“ (Nr. 19) fallen durch die hervorragende Zuordnung speziell des männlichen Sprechers auf. Dem könnte eine Anhäufung des dunkel klingenden Vokals /u/ zu Grunde liegen, der in den genannten Sätzen stets als langer Vokal auftritt und hierdurch in seiner Prägnanz an Bedeutung gewinnt.

Mit dieser Hypothese lassen sich die übrigen Ausreißersätze jedoch nicht erklären. In den Sätzen Nr. 14, Nr. 30 und Nr. 32 (vgl. Anhang A) sind hell und dunkel klingende Vokale relativ gleichmäßig verteilt und die Anzahl der Zischlaute ist mäßig. Es könnten andere, der Verfasserin der vorliegenden Arbeit unbekanntere Faktoren mitgewirkt haben und zu Überlagerungen mit oben beschriebener Hypothese geführt haben.

Demzufolge bleibt ungeklärt, ob und inwiefern die genannte Hypothese zutreffend ist. Da sie andererseits nicht ausgeschlossen werden kann, soll bei der Konstruktion neuer verbesserter Sätze auf eine Gleichverteilung von hell und dunkel klingenden Vokalen sowie von Zischlauten geachtet werden.

### 4.5 Fehlermöglichkeiten

Bei der Planung eines Experiments ist es wichtig, möglichst alle vermeidbaren Fehler auszuschließen. Zur Fehlerreduktion kamen die Vorversuche sowie eine gründliche Methodik zur Anwendung. Weiterhin sollten für alle Probanden die gleichen Versuchsbedingungen herrschen. Dies sollte mit Hilfe der exakt gleichen Instruktion sowie dem gleichen schallgedämpften Raum erfolgen. Andere, bei der Planung des Experiments nicht vorhersehbare oder nur schwer vermeidbare Fehlermöglichkeiten werden im folgenden beschrieben:

Bei der Beurteilung der Probandenantworten sind die interindividuellen Unterschiede in der Prosodiewahrnehmung zu berücksichtigen. Zugrunde liegende Faktoren könnten neben individuellen Hörgewohnheiten und Intellekt verschiedene Sprachkenntnisse der Versuchsteilnehmer oder eventuelle Erfahrungen in auditiver Arbeit sein. Letzteres konnte durch eine genaue Probandenauswahl jedoch weitestgehend ausgeschlossen werden. Fast alle Versuchsteilnehmer waren mit der Erstellerin der vorliegenden Arbeit persönlich bekannt. Effekte auf die Prosodiewahrnehmung wurden u. a. von folgenden Autoren untersucht: Nach [Artemov \(1956\)](#) ist die Perzeption der Intonation durch physikalische, grammatische, expressive sowie Bedeutungsfaktoren bedingt, die natürlicherweise personenabhängig sind. [Peppé et al. \(2000\)](#) beschreiben einen Effekt verursacht durch den Bildungsstand der Testpersonen, während [Magne et al. \(2006\)](#) eine Überlegenheit musikalischer Kinder bei der Prosodiewahrnehmung feststellten. Diese übertrafen unmusikalische Kinder bei der Perzeption von Grundfrequenzen.

Ebenfalls zu bedenken bleibt die für den Test nötige Änderung der Einstellung der Probanden. Während in der alltäglichen Kommunikation in erster Linie auf den Sinn einer Äußerung geachtet wird, sollte dieser bei Durchführung des Tests vernachlässigt werden. Vielmehr sollten prosodische Muster erkannt werden. Einige Autoren beschreiben diesen Umstand als eine Umwandlung einer funktionsorientierten Wahrnehmung im täglichen Sprachgebrauch zu einer signalorientierten Perzeption bei Durchführung des Tests. Somit wurde eine spezielle Abhöreinstellung abverlangt, die mit der alltäglichen Informationsverarbeitung nicht zu vergleichen ist.

Schwierigkeiten bereitet haben könnte auch der im folgenden erläuterte Umstand: Es besteht die Möglichkeit, dass Sätze aufgrund ihrer kontextunabhängigen separaten Darbietung eventuell nicht sicher beurteilt werden konnten. In der echten kommunikativen Situation wäre eine solche Äußerung dagegen durchaus eindeutig, weil hier der Perzipient auch auf nichtakustische Informationen zurückgreift ([Stock, 1980](#)). [Liu und Kewley-Port \(2007\)](#) untersuchten Einflüsse auf die Erkennung von Vokalformanten an Hörgeschädigten. Sie fanden ebenfalls heraus, dass die Erkennungsschwelle deutlich mit zunehmendem linguistischen Kontext (isolierte Vokale, Silben, Sätze) stieg. Dies zeigt, dass Testmaterial, bestehend aus vollständigen Sätzen, eine weitestgehend realitätsnahe Situation erzeugt, diese jedoch natürlicherweise nie vollständig erreichen kann.

Generell kann aus Wahrnehmungsvorgängen unter Laborbedingungen nicht ohne weiteres auf Prozesse der Alltagswahrnehmung geschlossen werden. Andererseits besteht der Zweck des in dieser Arbeit evaluierten Prosodietests vor allem darin,

den Test per se zu verifizieren und daraufhin die Leistungen Normalhörender mit denen von CI-Trägern zu vergleichen. Unterschiedliche Signalkodierungsstrategien der Cochleaimplantate könnten auf ihre Fähigkeit der Prosodieentschlüsselung untersucht werden.

Ein wesentliche Fehlermöglichkeit stellt die Beeinflussung der Versuchspersonen durch die Instruktion dar. Aus statistischen Gründen ist die Anzahl der jeweiligen Sprecher bzw. Kategorien pro Liste stets dieselbe. Auch alle Satzvarianten kommen im Test gleich oft vor. Die daraus entstehenden Probleme wurden bereits ausführlich in Abschnitt 2.7 erläutert. Änderungen im Verhalten einiger Probanden im Verlauf des Tests können daher nicht sicher ausgeschlossen werden.

Auch könnten vorausgehende Schallereignisse auf das Antwortverhalten der Probanden Einfluss genommen haben. Es wäre denkbar, dass jemand bei einer schwierigen Hörkondition innerhalb einer Liste zuerst beispielsweise die Frau mit dem Kind verwechselt und dann bei einem besonders eindeutigen Satz seinen Irrtum bemerkt und sein Verhalten ändert. Es ist jedoch unwahrscheinlich, dass solche Einflüsse eine große Wirkung auf die Versuchsergebnisse haben, zumal die Reihenfolge für die Items und die Listen für jeden Probanden eine andere war (s. Abschnitt 2.4.1 und 2.5.1). Durch die entsprechende Kombination von Items sowie Listen sollten diese Lernprozesse so auf ein Minimum reduziert werden.

Weiterhin gibt es bei allen Experimenten stets unvermeidliche Einflüsse, die aber so gering wie möglich gehalten werden sollten. So zweifelten bei den lautereren Hörkonditionen die meisten Probanden natürlicherweise an den abgegebenen Urteilen. Manche gaben jedoch ein Gefühl der Frustration an, trotz regelrechter Bemühungen das Item nicht sicher zu erkennen. Diese Gemütszustände könnten somit eventuell zu Motivationsmängeln und Konzentrationsschwächen geführt haben. Ebenso wäre es möglich, dass der von den meisten Versuchsteilnehmern als anstrengender beschriebene Kategorientest in einem schnelleren Energieverbrauch und so ebenfalls in einer nachlassenden Konzentration resultiert. Andere Einflüsse in Bezug auf die Konzentration könnten in der ungewohnten Umgebung in Form der schallgedämpften Hörkabine liegen. Einige Versuchspersonen empfanden die Stille als sehr angenehm, während diese von anderen als eher beängstigend beschrieben wurde. Um Konzentrationsschwächen und Motivationsmängel möglichst zu reduzieren, konnten die Probanden jederzeit Pausen in Anspruch nehmen.

## 5 Ausblick

Die Evaluation des Tests zeigt, dass die Prosodieperzeption Normalhörender mit dem vorliegenden Instrumentarium umfassend getestet werden kann. Im Hinblick auf eine zukünftige Anwendung des Prosodietests mit Hörhilfeträgern ergeben sich viele weitere Fragestellungen. Einige der wichtigsten werden im folgenden genannt:

- Inwieweit weist der Test das Potenzial auf, Unterschiede in der Prosodiewahrnehmung von Hörgeschädigten oder Hörhilfeträgern bestimmen zu können?
- Ist der Test für Vergleiche und Optimierungen unterschiedlicher moderner Sprachsignalverarbeitungsstrategien adäquat?
- Ist die Hypothese der besseren Wahrnehmbarkeit von Grundfrequenzverläufen und somit prosodischer Aspekte mit Hilfe der neuen CSSS-Strategie anhand des Tests verifizierbar?
- Inwieweit liegt eine Beeinflussung durch die Satzausreißer vor und wie sollte mit diesen verfahren werden?
- Welchen Einfluss hat die Probandeninstruktion?

Der Prosodietest wurde ursprünglich entwickelt, um besonders die ersten drei Fragestellungen damit untersuchen zu können. Bei Durchführung und Auswertung des Tests ergaben sich u. a. die beiden letztgenannten Fragen, deren Untersuchung zu einer Optimierung des Testinstrumentariums führen könnte.

Im Hinblick auf die Ausreißersätze ergeben sich mehrere Möglichkeiten, mit diesen weiter zu verfahren. Zum gegenwärtigen Zeitpunkt stehen insgesamt 36 verschiedene Sätze zur Verfügung, die alle in jeder Liste jeweils einmal vertreten sind. Man könnte nun einfach einige der ungeeignetsten Sätze streichen und hätte ein leicht reduziertes Testmaterial ausschließlich geeigneter Sätze vorliegen. Auf diese Weise könnte auf die Entwicklung und Integrierung neuer Testsätze verzichtet werden, die in einem weiteren Schritt ebenfalls erst an Normalhörenden getestet werden müssten. Zur Beibehaltung eines homogen strukturierten Testmaterials wäre eine Anzahl von insgesamt 27 Sätzen (Dreierpotenz) erstrebenswert. Die Sprecher Mann, Frau,

Kind sowie die Kategorien Aussage, Frage und Ausruf wären so pro Liste gleich oft vertreten. Es müsste auf neun Sätze verzichtet werden. Die Erstellerin der vorliegenden Arbeit schlägt vor, in diesem Fall die Ausreißer der Testart Kategorie (Satz-Nrn. 3, 21, 22, 25, 15, 6, 10, 23) sowie den stärksten Ausreißer der Sprecher (Satz-Nr. 14) zu streichen (vgl. Abschnitt 4.4).

Andererseits könnte man auch folgendermaßen argumentieren: Da nach Auswertung des Tests genau bekannt ist, welche Sätze schlechter (oder in einem Fall besser) erkannt werden und das Verhalten dieser Ausreißersätze umfassend untersucht wurde, könnte man die betreffenden Sätze mit diesem Wissen beibehalten. Bei Folgeanwendungen des Tests müsste dieser Umstand jedoch Berücksichtigung finden. Dieser Möglichkeit kommt zugute, dass es sich bei den oben genannten Sätzen nur im weitesten Sinn um Satzausreißer und eher um geringgradige Abweichungen handelt.

Außerdem besteht die Möglichkeit, die Ausreißersätze durch neu entworfene verbesserte Sätze zu ersetzen. In Anhang D wird eine Liste mit einigen neuen Sätzen vorgestellt, die eventuell in den Test aufgenommen werden könnten. Einer Anwendung an Patienten mit Hörhilfen würde jedoch eine vorherige Untersuchung mit normalhörenden Probanden vorausgehen.

In Abschnitt 2.7 wurde bereits das Problem des gleichhäufigen Vorkommens der Satzvarianten innerhalb aller Listen sowie der gleichen Anzahl der Sprecher- bzw. Kategorienvarianten pro Liste beschrieben. Jeder Satz wurde in den drei Varianten jeweils zweimal pro Testart (inkl. Übungsdurchlauf) verwendet. In jeder Liste kamen je zwölf der unterschiedlichen Sprecher (Mann, Frau und Kind) sowie je zwölf der verschiedenen Kategorien (Aussage, Frage und Ausruf) vor. Hierüber wurden die Testpersonen aus den in Abschnitt 2.7 erläuterten Gründen nicht aufgeklärt. Die hieraus entstehende Problematik liegt darin, dass das gleiche Vorkommen von Satzvarianten und Sprecher- bzw. Kategorienvarianten von den Probanden bemerkt worden sein könnte, worauf diese ihr Antwortverhalten entsprechend geändert haben könnten.

Man könnte nun andererseits im Vorfeld die Versuchsteilnehmer aufklären und so für alle gleiche Bedingungen schaffen. Allerdings bestünde dann die Gefahr, dass das Antwortverhalten dadurch noch stärker beeinflusst wird, da manche das gleichhäufige Vorkommen wahrscheinlich gar nicht bemerkten und das interindividuelle Verhalten der Probanden naturgemäß weit gestreut ist.

Weiterhin könnte man mit der Argumentation, dass nach Testauswertung nun die Beeinflussung auf die Testpersonen einschätzbar ist, die Instruktion so belassen. Man hätte so vergleichbare Ergebnisse für Folgetests.

Als dritte Möglichkeit könnte nur jeweils eine Variante pro Satz für jede Testart dargeboten werden, um zumindest einen der Einflussfaktoren ausschließen zu können. Allerdings würde dadurch das Testmaterial auf ein Sechstel seiner ursprünglichen Größe schrumpfen und eine entsprechende Anzahl neuer Sätze müsste generiert werden. Diese müssten wiederum an Normalhörenden auf ihre Eignung untersucht werden.





## 6 Zusammenfassung

Die Wahrnehmung prosodischer Merkmale geht über die reine Semantik des Sprachverstehens hinaus und beeinflusst maßgeblich die Lebensqualität von Personen mit Hörstörungen oder Hörhilfen, die in der Prosodieperzeption eingeschränkt sind. Zur Untersuchung der Wahrnehmung prosodischer Aspekte liegt derzeit im deutschsprachigen Raum kein adäquates Testinstrumentarium vor. Aufbauend auf der Vorgängerarbeit von Bauer (2005) wurde der bereits bestehende Prosodietest weiterentwickelt und an einer größeren Anzahl Normalhörender, die bei späteren Tests mit Hörhilfeträgern als Vergleichsgruppe fungieren sollen, auf sein Eignungspotential untersucht.

Der Prosodietest besteht aus zwei unterschiedlichen Testarten (Sprecher- und Kategorientest), die verschiedene prosodische Merkmale abdecken. Beim Sprechertest sollen die drei Sprecher Mann, Frau und Kind erkannt werden, deren Unterscheidbarkeit auf den entsprechenden Tonhöhen (Grundfrequenzen) beruht. Der Kategorientest zielt auf die Erkennung von Grundfrequenzverläufen sowie Akzenten ab, die typischerweise bei den dargebotenen Kategorien Aussage, Frage und Ausruf unterschiedlich ausgeprägt sind. Das Testmaterial wurde in Form von fünf Listen pro Testart mit jeweils 36 Sätzen präsentiert, wobei die Sätze für die jeweiligen Kategorien sowie Sprecher syntaktisch und grammatisch identisch sind. Auf diese Weise wird ausschließlich das Prosodieverständnis abgefragt. Es wurden 30 junge Erwachsene im Alter zwischen 20 und 30 Jahren ausgewählt und deren Normalhörigkeit mittels eines Tonaudiogramms festgestellt. Durch Überlagerung der Testsätze mit einem Störgeräusch in verschiedenen Lautstärken wurden unterschiedliche Schwierigkeitsstufen geschaffen. Zur Anwendung kam ein CCITT-Rauschen, das nach Vorversuchen zur Einschätzung der Prosodieperzeption Normalhörender, in den sechs Hörkonditionen 0 dB SNR, -6 dB SNR, -12 dB SNR, -18 dB SNR und -24 dB SNR verwendet wurde. Als leichter Übungsdurchlauf wurde zusätzlich die Hörkondition +6 dB SNR dargeboten. Um möglichst gleiche Bedingungen für alle Versuchsteilnehmer zu schaffen und unerwünschte Einflussfaktoren so gering wie möglich zu halten, wurde bei der Konzeption des Tests das Verfahren der permutierten Blockrandomisierung angewendet. Des Weiteren erhielten alle Probanden die gleiche schriftliche Instruktion.

Erwartungsgemäß lagen die Ergebnisse der Probanden im Bereich zwischen der 33,3%-Marke des reinen Ratens und der 100%-Grenze der maximalen Leistung. Die Ergebnisse zeigten bei der Darstellung der Gesamtdaten relativ geringe Streuungen der logistischen Regressionen der einzelnen Probanden sowie deutliche Wendepunkte. Dies spricht für eine hohe Sensitivität des Tests. Der Wendepunkt lag beim Sprechertest bei  $-14,72$  dB SNR, beim Kategorientest bei  $-11,65$  dB SNR, wodurch dieser Test als schwieriger einzustufen ist. Die Anzahl der Verwechslungen von Urteilen stieg vom leichten bis in den schwierigen Hörbereich kontinuierlich an, wobei eine wachsende Ungleichmäßigkeit in der Verteilung der Urteile auffiel. Bei totalem Raten der Versuchsteilnehmer existierte bei den Sprechern für den Mann eine Ratewahrscheinlichkeit von 50,3 %, für die Frau von 31,9 % und für das Kind von nur 17,8 %. Im Kategorientest wurden für die Aussage 45,7 %, für die Frage 23,1 % und für den Ausruf 31,2 % ermittelt. Bei der satzweisen Auswertung des Tests fielen insbesondere beim Kategorientest einige „Satzausreißer“ auf, die geringgradige Abweichungen vom Gesamtdatensatz darstellten und als Interpretationsfehler angesehen werden müssen. Beim Kategorientest stellten vor allem die Sätze Nr. 3, Nr. 21, Nr. 22 und Nr. 25 Negativausreißer sowie Nr. 15 einen extrem gut erkannten Satz dar, während beim Sprechertest die Sätze Nr. 5, Nr. 14 und Nr. 30 etwas weniger gute Erkennungsraten aufwiesen.

Durch die Evaluation des Prosodietests mit normalhörenden Probanden hat sich gezeigt, dass dieser Test das Potential besitzt, brauchbare und umfassende Informationen über die Erfassbarkeit der Prosodie zu erzielen. Um den in dieser Arbeit weiterentwickelten und evaluierten Test in einen Test für die klinische Praxis zu überführen, sind nur noch wenige Arbeitsschritte nötig. Es müssten lediglich die weiter oben genannten „Satzausreißer“ unter Beachtung der in Kapitel 5 aufgeführten Gesichtspunkte entfernt oder durch neue Sätze ersetzt werden (s. Anhang D), bevor der Test zum Einsatz gebracht werden kann.

# Anhang



## A Testsätze

1. Da ist das Rathaus
2. Jetzt gibt es schon Eis
3. Schön haben wir es hier
4. Der Computer ist neu
5. Mit dem Taschenrechner kannst du das
6. Der Fotoapparat funktioniert wieder
7. Am Fenster sitzt eine Katze
8. Der Hund bellt
9. Susi kocht das Essen
10. In der Schule lernt man Mathematik
11. Der Großvater ist krank
12. Der Sessel ist bequem
13. Gleich fährt der Zug
14. Die Tagung war erfolgreich
15. Das Konzert ist super
16. Morgen soll die Sonne scheinen
17. Am Montag findet ein Marathon statt
18. Du hast Besuch
19. Das Buch ist spannend
20. Der Schnellzug geht um eins

## A Testsätze

21. Der Maler ist wirklich begabt
22. Klettern kann gefährlich sein
23. Der Koffer ist schwer
24. Das Kreuzworträtsel ist einfach
25. Berlin ist eine große Stadt
26. Draußen gibt es viele Mücken
27. Am Strand liegen viele Muscheln
28. Der Asphalt ist heiß geworden
29. Die Katze kann kläglich schreien
30. Benzin ist wieder teurer geworden
31. Kinoplakate sind sehr groß
32. Die Nachbarn sind laut
33. Der Fernseher ist kaputt
34. Die Reise war anstrengend
35. Die Vermieter sind freundlich
36. Der Hund beißt

## B Kontingenztabelle

Auf den folgenden drei Seiten sind die Kontingenztabelle der Sprecher und Kategorien abgebildet. Jeweils eine Tabelle pro Testart (vgl. B.1 und B.7) zeigt die Anzahl und Verteilung der Probandenurteile für alle SNRs kombiniert und soll einen Überblick verschaffen, während in fünf weiteren Kontingenztabelle die Urteile für jeden SNR separat dargestellt sind.

In der jeweils linken Spalte sind die dargebotenen Sprecher M, W und K bzw. die dargebotenen Kategorien A, F und R aufgetragen. Die obere Zeile gibt die abgegebenen Urteile an. Im Gegensatz zu dieser tabellentypischen Schreibweise sind die dargebotenen Sprecher bzw. Kategorien des Blasendiagramms 3.4 (s. Seite 44) auf der rechten Seite und die Urteile unten aufgezeichnet. So entspricht beispielsweise die untere linke Blase des Diagramms nicht dem linken unteren, sondern dem linken oberen Feld der zugehörigen Kontingenztabelle. Diese unterschiedliche Positionierung von Blasen und zugehörigen Tabellenwerten muss bei Betrachtung der verschiedenen Abbildungen Beachtung finden.

Zur Verdeutlichung der Anzahl der korrekten Urteile müssen also die Diagonalen von links oben nach rechts unten betrachtet werden. Verwechslungen werden bei zeilenweiser Betrachtung erkennbar. Die Urteilshäufigkeit geht aus der jeweils unteren Summenzeile hervor.

Die jeweils rechte Spalte gibt zur Verdeutlichung die Anzahl der dargebotenen Testsätze an. Bei den Tabellen der kombinierten SNRs und 5400 dargebotenen Sätzen pro Testart (rechtes unteres Feld) wurden dementsprechend jeweils 1800 der drei verschiedenen Sprecher bzw. Kategorien angeboten (s. rechte Spalte). In den Kontingenztabelle, in denen die Werte für jeden SNR separat gezeigt werden, sind demzufolge 1080 Sätze pro Testart und 360 Sätze pro Sprecher bzw. Kategorie in der rechten Spalte abgebildet. Die untere Summenzeile stellt die Häufigkeit der abgegebenen Urteile dar.

Tabelle B.1: Kontingenztafel der Sprecher, alle SNRs kombiniert

| Sprecher | Urteil |      |      | $\Sigma$ |
|----------|--------|------|------|----------|
|          | M      | W    | K    |          |
| M        | 1613   | 110  | 77   | 1800     |
| W        | 161    | 1418 | 221  | 1800     |
| K        | 505    | 377  | 918  | 1800     |
| $\Sigma$ | 2279   | 1905 | 1216 | 5400     |

Tabelle B.2: Kontingenztafel der Sprecher bei 0 dB SNR

| SNR: 0 dB<br>Sprecher | Urteil |     |     | $\Sigma$ |
|-----------------------|--------|-----|-----|----------|
|                       | M      | W   | K   |          |
| M                     | 360    | 0   | 0   | 360      |
| W                     | 0      | 355 | 5   | 360      |
| K                     | 0      | 0   | 360 | 360      |
| $\Sigma$              | 360    | 355 | 365 | 1080     |

Tabelle B.3: Kontingenztafel der Sprecher bei -6 dB SNR

| SNR: -6 dB<br>Sprecher | Urteil |     |     | $\Sigma$ |
|------------------------|--------|-----|-----|----------|
|                        | M      | W   | K   |          |
| M                      | 355    | 4   | 1   | 360      |
| W                      | 1      | 319 | 40  | 360      |
| K                      | 6      | 44  | 310 | 360      |
| $\Sigma$               | 362    | 367 | 351 | 1080     |

Tabelle B.4: Kontingenztafel der Sprecher bei -12 dB SNR

| SNR: -12 dB<br>Sprecher | Urteil |     |     | $\Sigma$ |
|-------------------------|--------|-----|-----|----------|
|                         | M      | W   | K   |          |
| M                       | 343    | 11  | 6   | 360      |
| W                       | 24     | 280 | 56  | 360      |
| K                       | 109    | 116 | 135 | 360      |
| $\Sigma$                | 476    | 407 | 197 | 1080     |



Tabelle B.5: Kontingenztabelle der Sprecher bei -18 dB SNR

| SNR: -18 dB<br>Sprecher | Urteil |     |     | $\Sigma$ |
|-------------------------|--------|-----|-----|----------|
|                         | M      | W   | K   |          |
| M                       | 308    | 38  | 14  | 360      |
| W                       | 52     | 243 | 65  | 360      |
| K                       | 188    | 118 | 54  | 360      |
| $\Sigma$                | 548    | 399 | 133 | 1080     |

Tabelle B.6: Kontingenztabelle der Sprecher bei -24 dB SNR

| SNR: -24 dB<br>Sprecher | Urteil |     |     | $\Sigma$ |
|-------------------------|--------|-----|-----|----------|
|                         | M      | W   | K   |          |
| M                       | 247    | 57  | 56  | 360      |
| W                       | 84     | 221 | 55  | 360      |
| K                       | 202    | 99  | 59  | 360      |
| $\Sigma$                | 533    | 377 | 170 | 1080     |

Tabelle B.7: Kontingenztabelle der Kategorien, alle SNRs kombiniert

| Kategorie | Urteil |      |      | $\Sigma$ |
|-----------|--------|------|------|----------|
|           | A      | F    | R    |          |
| A         | 1310   | 192  | 298  | 1800     |
| F         | 359    | 1155 | 286  | 1800     |
| R         | 499    | 240  | 1061 | 1800     |
| $\Sigma$  | 2168   | 1587 | 1645 | 5400     |

Tabelle B.8: Kontingenztabelle der Kategorien bei 0 dB SNR

| SNR: 0 dB<br>Kategorie | Urteil |     |     | $\Sigma$ |
|------------------------|--------|-----|-----|----------|
|                        | A      | F   | R   |          |
| A                      | 332    | 0   | 28  | 360      |
| F                      | 1      | 350 | 9   | 360      |
| R                      | 11     | 2   | 347 | 360      |
| $\Sigma$               | 344    | 352 | 384 | 1080     |

Tabelle B.9: Kontingenztabelle der Kategorien bei  $-6$  dB SNR

| SNR: -6 dB<br>Kategorie | Urteil |     |     | $\Sigma$ |
|-------------------------|--------|-----|-----|----------|
|                         | A      | F   | R   |          |
| A                       | 313    | 15  | 32  | 360      |
| F                       | 13     | 336 | 11  | 360      |
| R                       | 56     | 24  | 280 | 360      |
| $\Sigma$                | 382    | 375 | 323 | 1080     |

Tabelle B.10: Kontingenztabelle der Kategorien bei  $-12$  dB SNR

| SNR: -12 dB<br>Kategorie | Urteil |     |     | $\Sigma$ |
|--------------------------|--------|-----|-----|----------|
|                          | A      | F   | R   |          |
| A                        | 278    | 29  | 53  | 360      |
| F                        | 56     | 256 | 48  | 360      |
| R                        | 104    | 61  | 195 | 360      |
| $\Sigma$                 | 438    | 346 | 296 | 1080     |

Tabelle B.11: Kontingenztabelle der Kategorien bei  $-18$  dB SNR

| SNR: -18 dB<br>Kategorie | Urteil |     |     | $\Sigma$ |
|--------------------------|--------|-----|-----|----------|
|                          | A      | F   | R   |          |
| A                        | 207    | 68  | 85  | 360      |
| F                        | 134    | 127 | 99  | 360      |
| R                        | 161    | 77  | 122 | 360      |
| $\Sigma$                 | 502    | 272 | 306 | 1080     |

Tabelle B.12: Kontingenztabelle der Kategorien bei  $-24$  dB SNR

| SNR: -24 dB<br>Kategorie | Urteil |     |     | $\Sigma$ |
|--------------------------|--------|-----|-----|----------|
|                          | A      | F   | R   |          |
| A                        | 180    | 80  | 100 | 360      |
| F                        | 155    | 86  | 119 | 360      |
| R                        | 167    | 76  | 117 | 360      |
| $\Sigma$                 | 502    | 242 | 336 | 1080     |

## C Einzelsatzgraphiken

Die folgenden zwölf Seiten zeigen Graphiken zur Beurteilung der Eignung aller 36 Sätze. Auf den ersten sechs Seiten sind die Einzelsatzgraphiken für den Kategorien-test dargestellt, die Abbildungen für den Sprechertest befinden sich auf den darauf folgenden sechs Seiten. Auf jeder Seite sind jeweils sechs Sätze in fortlaufender Reihenfolge abgebildet, wobei jeder Satz eine Zeile einnimmt. Die Satznummern (rechts unten in der vierten Spalte) bezeichnen den jeweiligen Satz. Alle Testsätze sind nummeriert in Anhang A aufgeführt. Die Erzeugung der Funktionen wurde bereits in Abschnitt 2.9 beschrieben.

Die ersten drei Spalten differenzieren zwischen den unterschiedlichen dargebotenen Kategorien A, F, und R sowie den Sprechern M, W und K. Die stärker hervorgehobenen Graphen (Punkte sind durch Linien verbunden) stellen die Urteile der Versuchsteilnehmer für den betreffenden Satz dar, während die durchgezogenen dünneren Kurven die Probandenurteile für die Gesamtheit aller Sätze abbilden. Auf diese Weise wird ein Vergleich des jeweiligen Satzes mit dem kompletten Testsatzmaterial möglich und eventuelle Ausreißer sind erkennbar.

Während aus den drei ersten Spalten die Verteilung der Urteile hervorgeht, stellt die vierte Spalte ausschließlich die korrekten Antworten jedes Satzes dar. Die Graphen der Einzelsätze sind – wie in den anderen Spalten – hervorgehoben und durch Punkte verbindende kurze Linien dargestellt. Um einen direkten Vergleich der einzelnen Sätze mit dem Gesamtdatensatz zu ermöglichen, sind die Graphiken mit den logistischen Regressionen unterlegt abgebildet.

Auf der Abszisse sind die Hörkonditionen von 0 dB SNR bis –24 dB SNR aufgetragen, die Ordinate zeigt die relative Häufigkeit der abgegebenen Urteile an. Da die ersten drei Spalten die Verteilung der Urteile auf die drei jeweiligen Antwortmöglichkeiten darstellen, wird die gesamte Skala der relativen Häufigkeit zwischen 0,0 und 1,0 eingesetzt. Bei der vierten Spalte wird nur ein Teilbereich der Ordinate zwischen 0,33 und 1,0 verwendet (vgl. Abschnitt 2.9.2 und 2.9.1).

### C Einzelsatzgraphiken

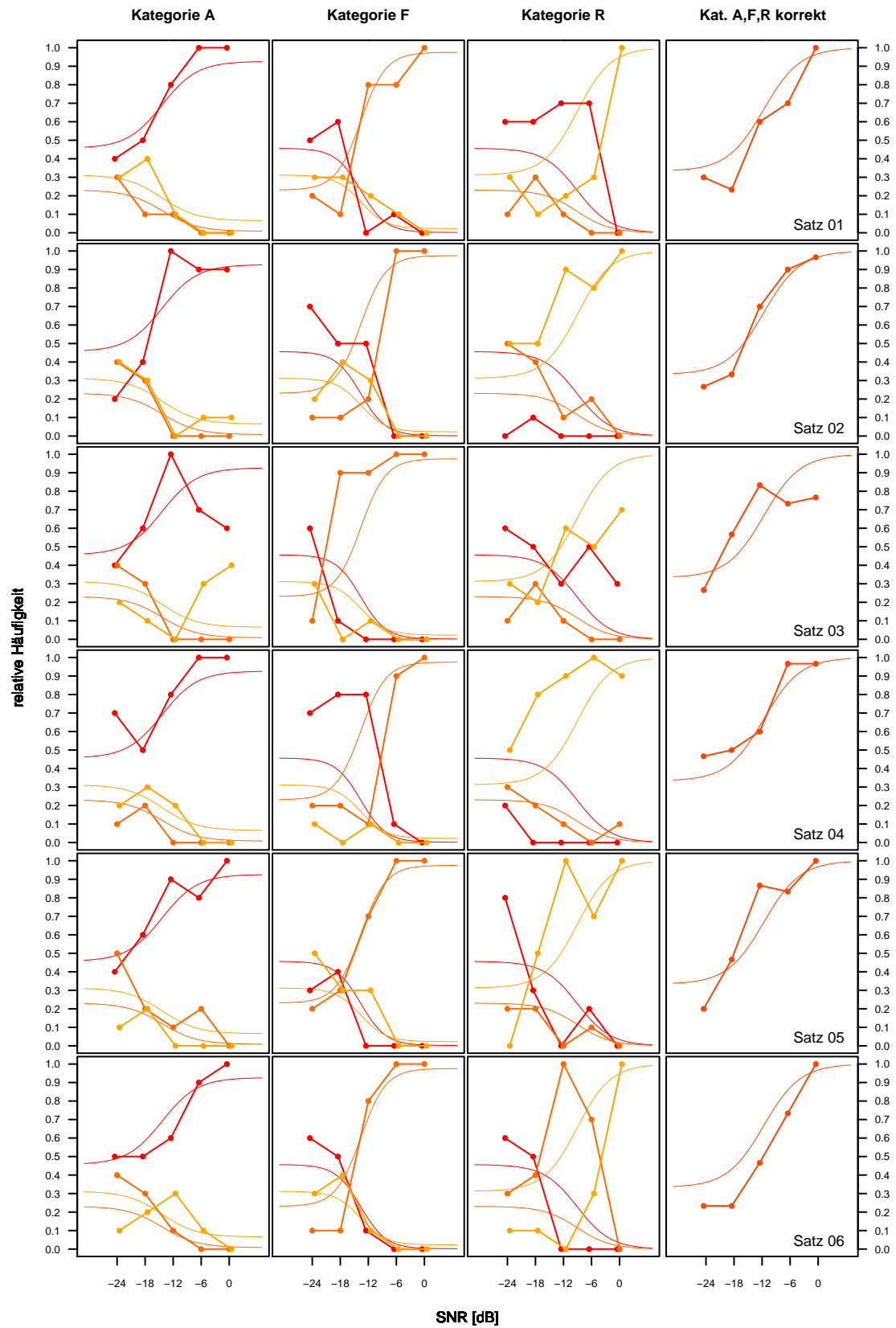


Abbildung C.1: Einzelsatzgraphiken der Testart Kategorie, Satz-Nr. 1–6

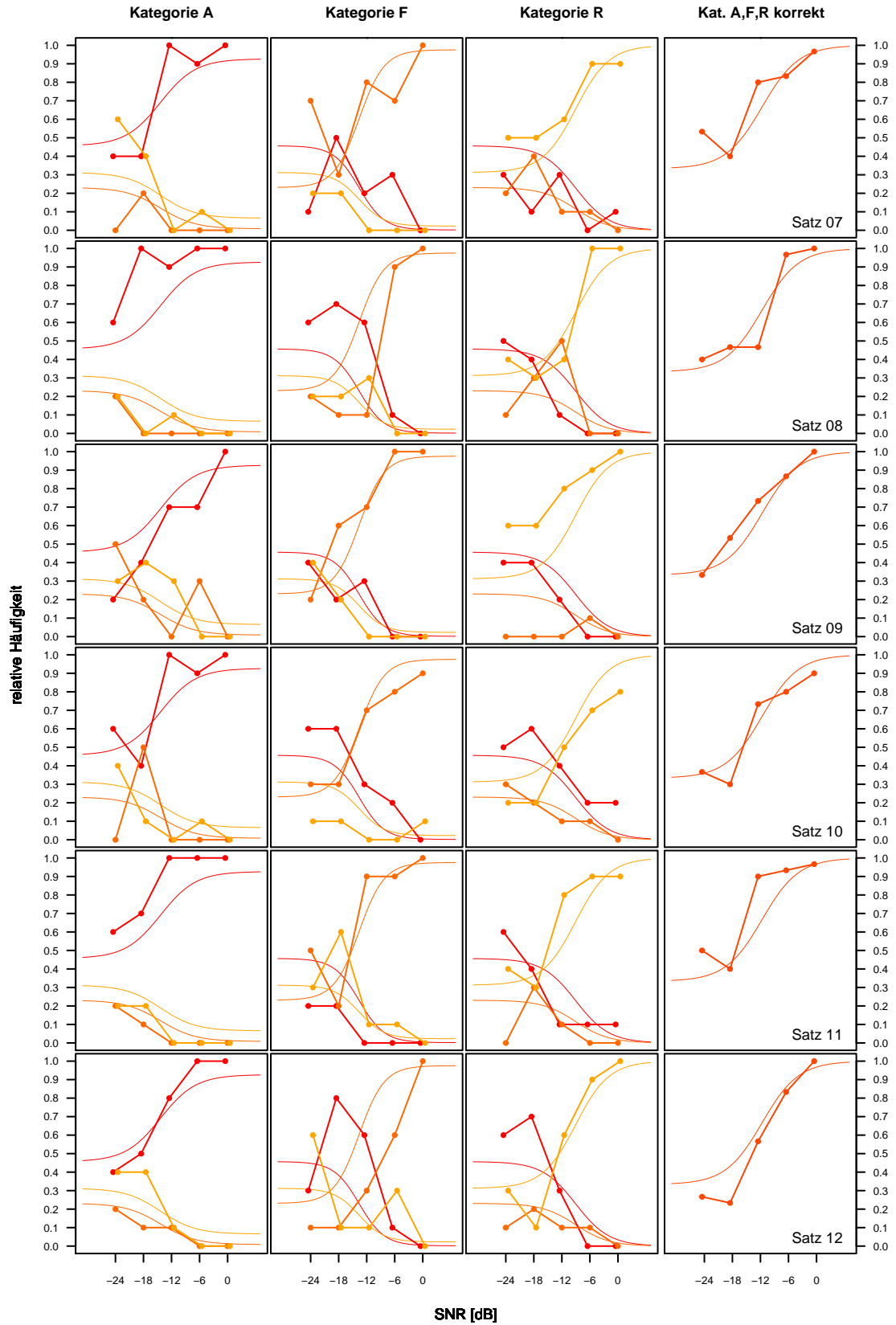


Abbildung C.2: Einzelsatzgraphiken der Testart Kategorie, Satz-Nr. 7–12

### C Einzelsatzgraphiken

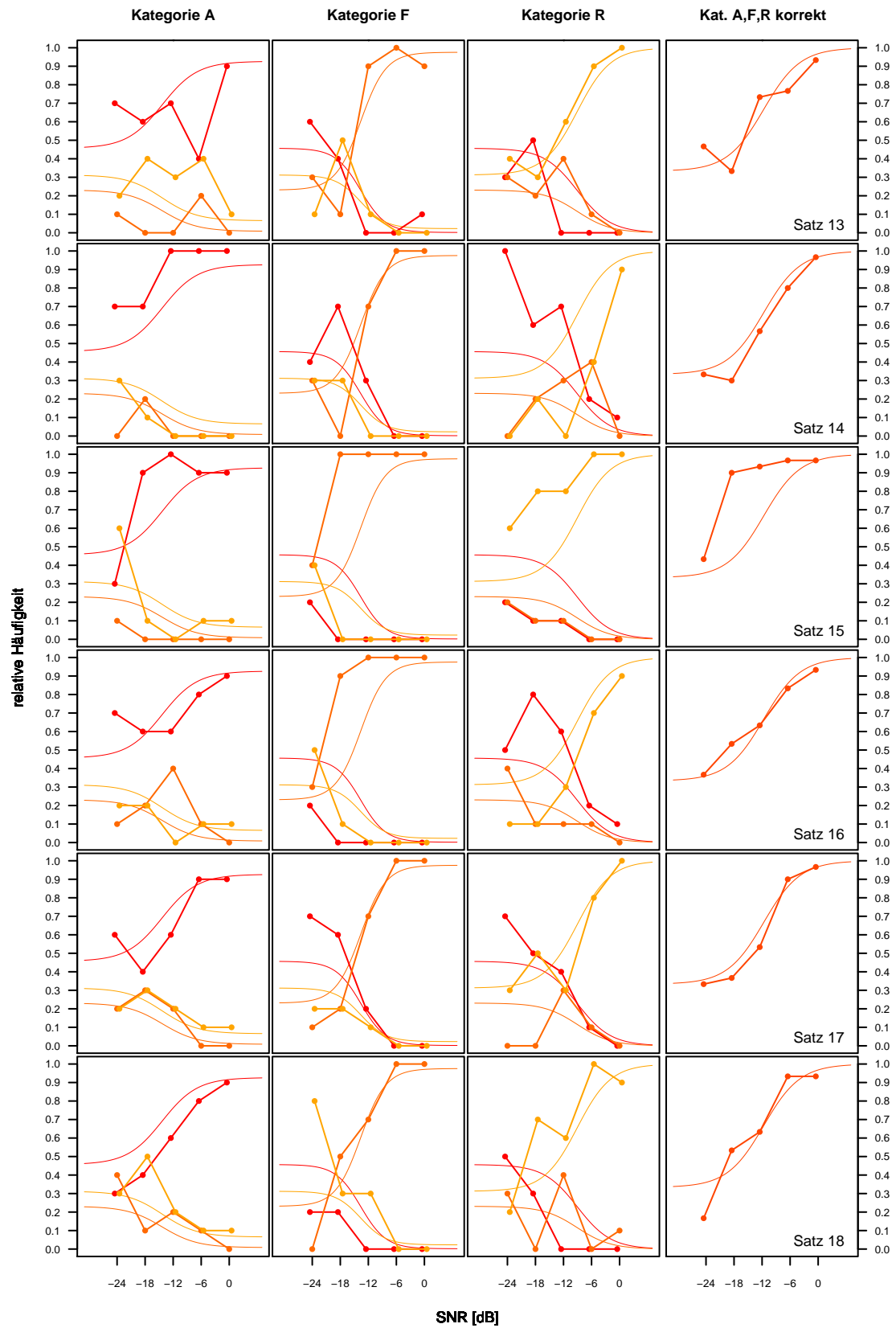


Abbildung C.3: Einzelsatzgraphiken der Testart Kategorie, Satz-Nr. 13–18

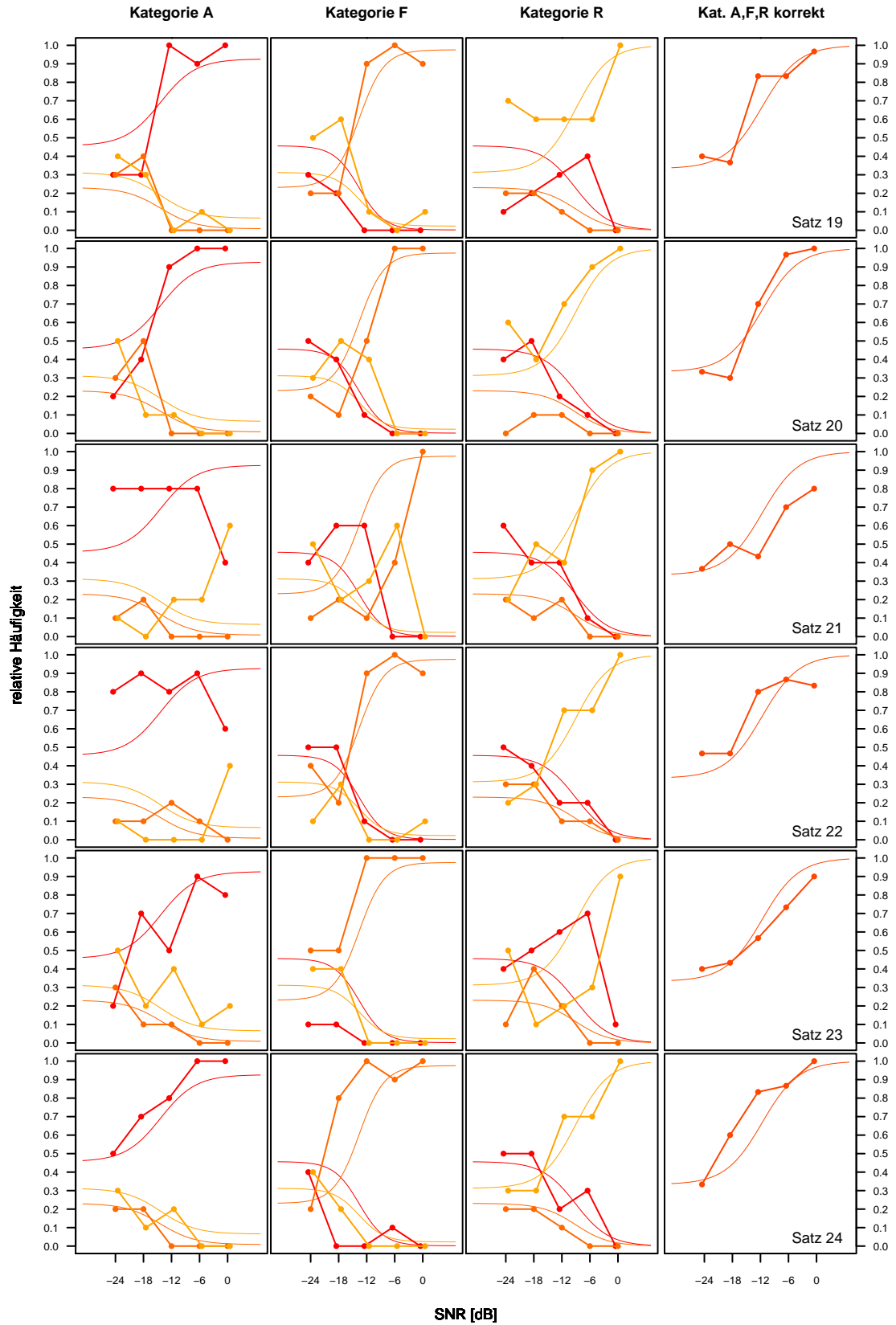


Abbildung C.4: Einzelsatzgraphiken der Testart Kategorie, Satz-Nr. 19–24

### C Einzelsatzgraphiken

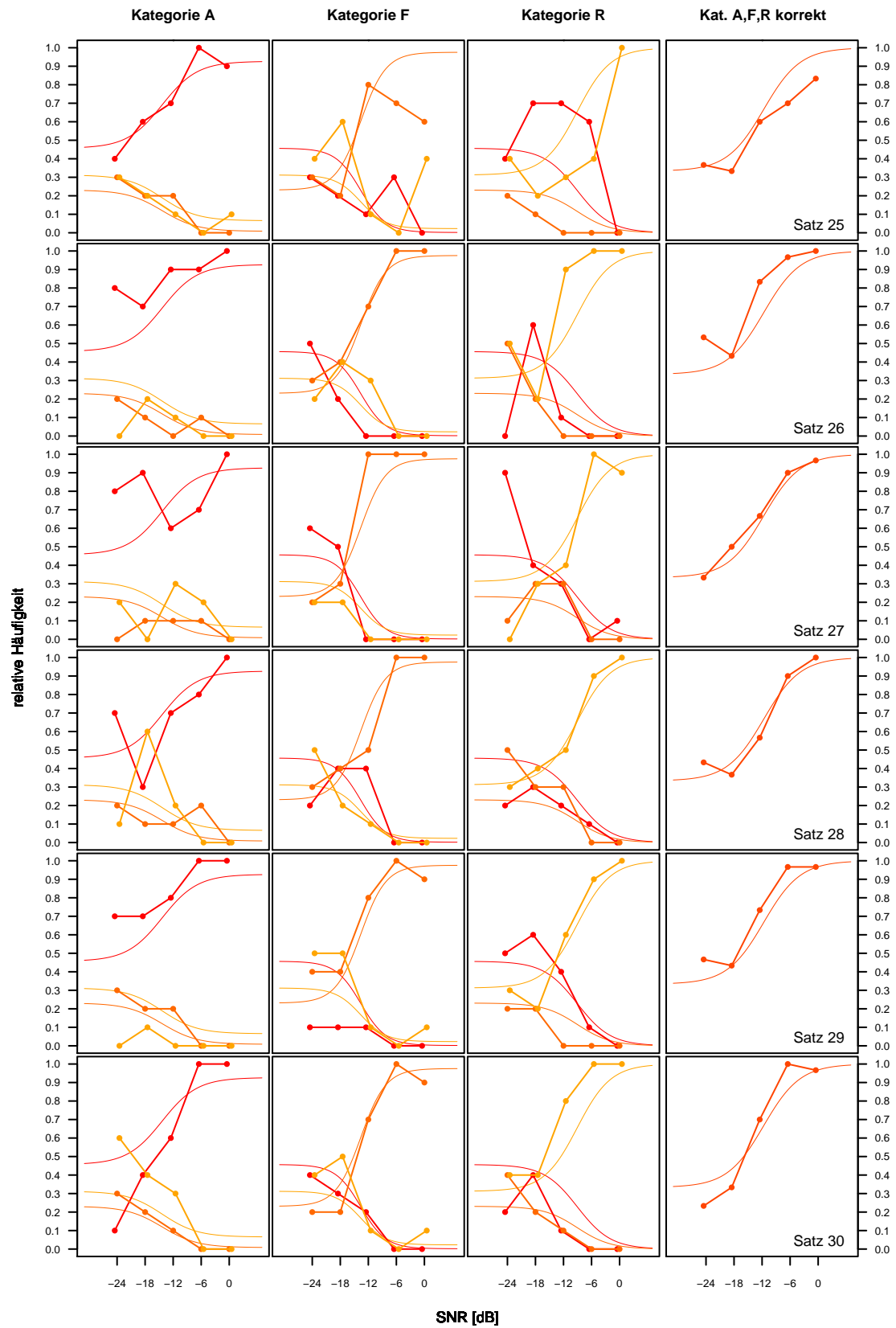


Abbildung C.5: Einzelsatzgraphiken der Testart Kategorie, Satz-Nr. 25–30



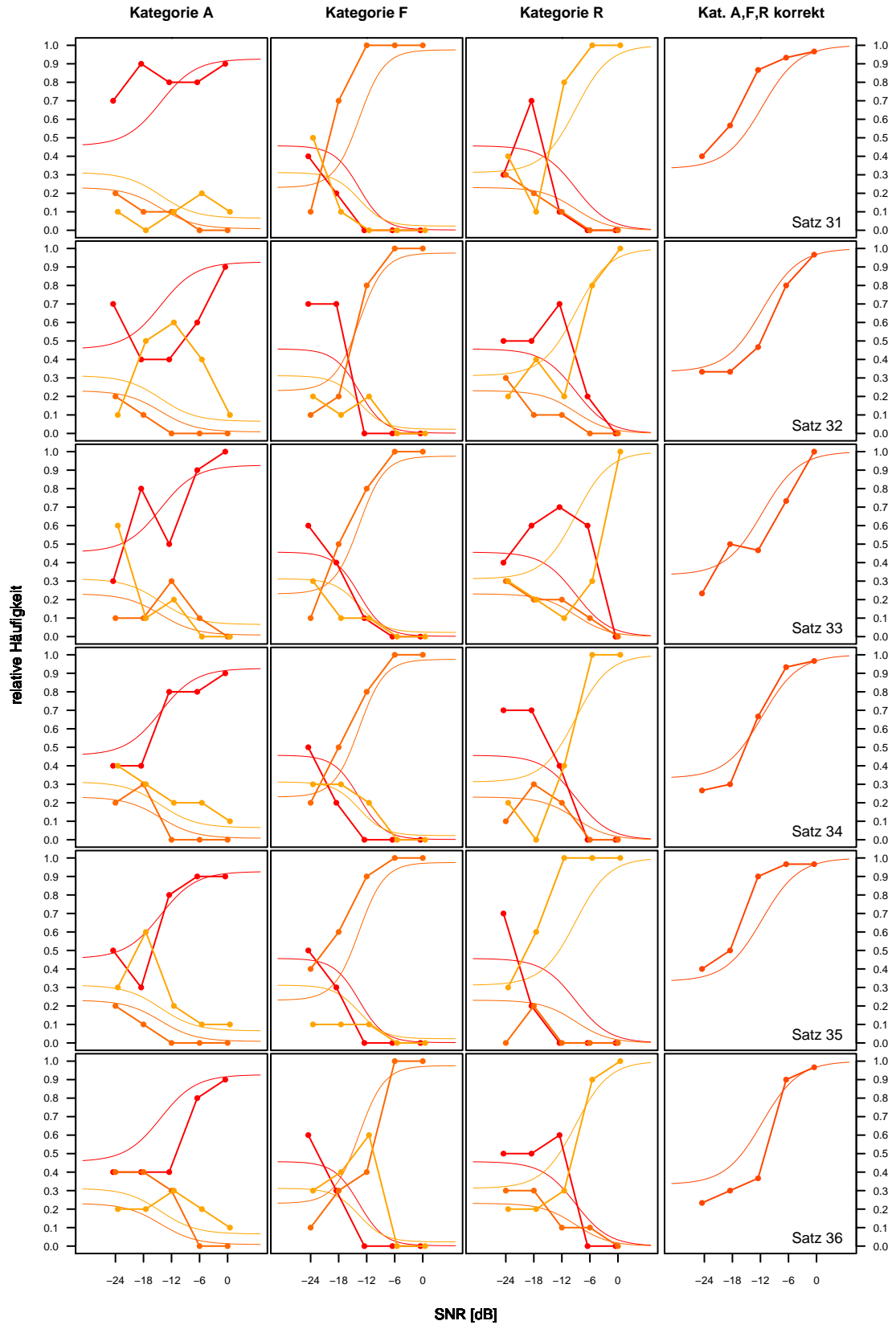


Abbildung C.6: Einzelsatzgraphiken der Testart Kategorie, Satz-Nr. 31–36

### C Einzelsatzgraphiken

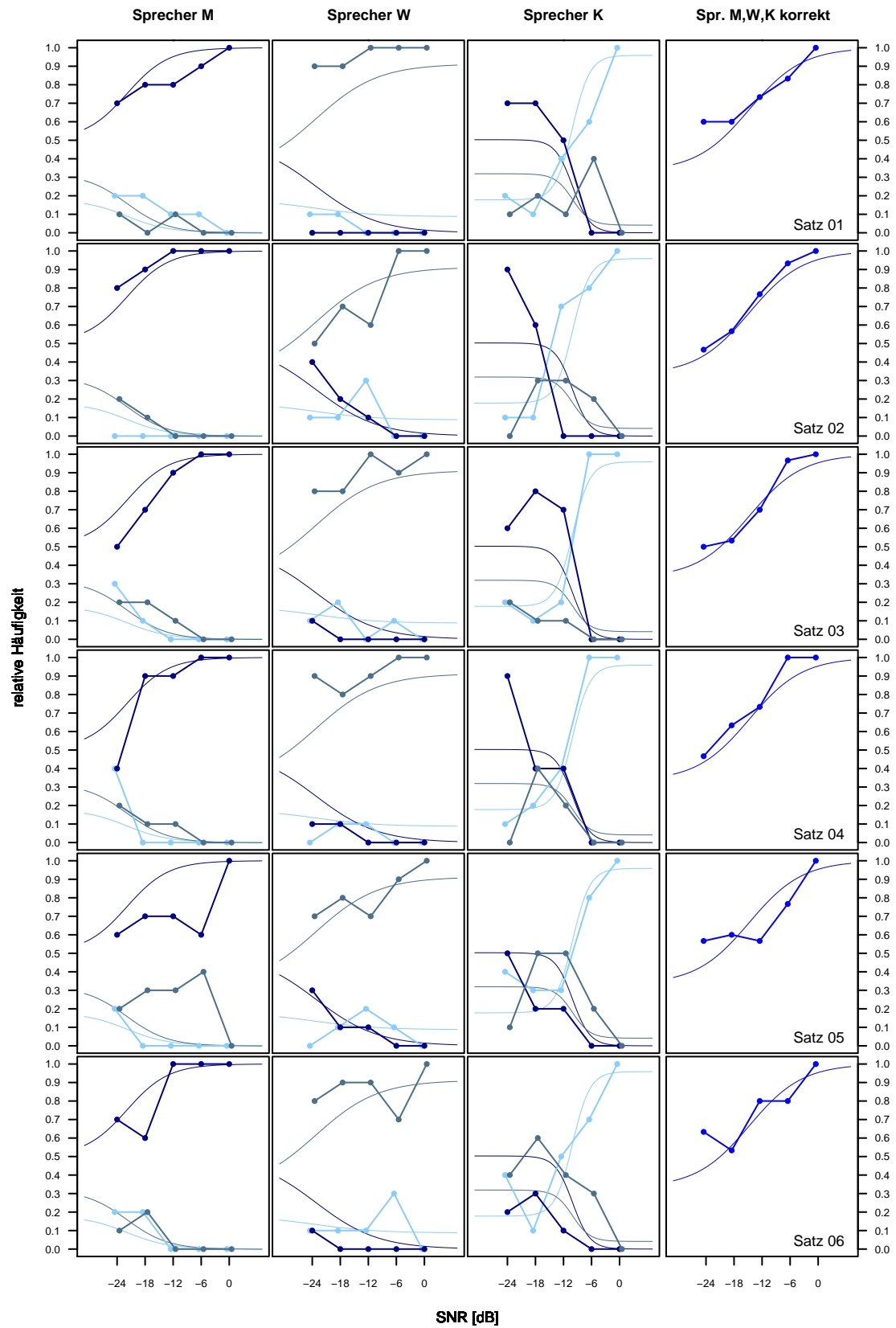


Abbildung C.7: Einzelsatzgraphiken der Testart Sprecher, Satz-Nr. 1–6

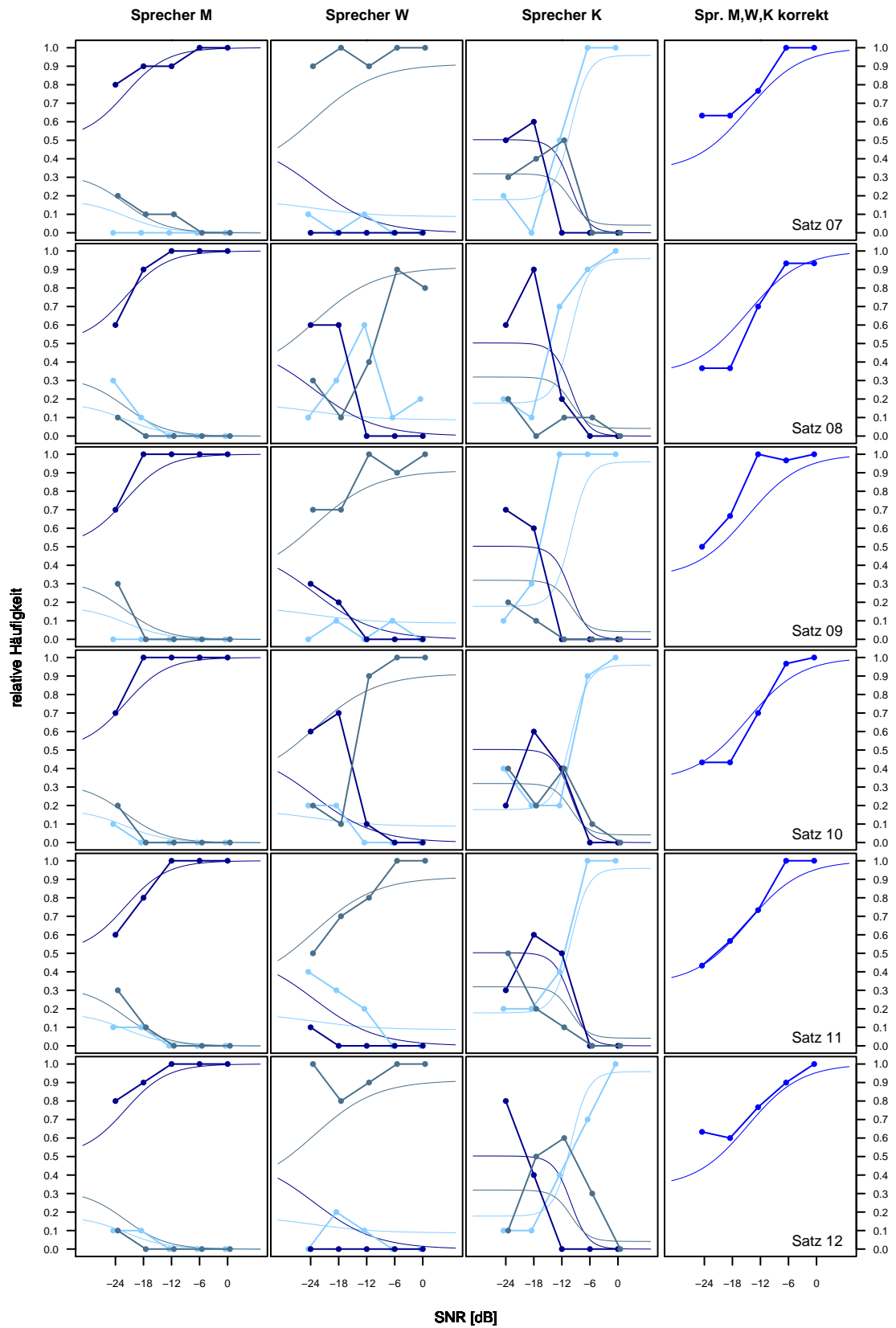


Abbildung C.8: Einzelsatzgraphiken der Testart Sprecher, Satz-Nr. 7–12

### C Einzelsatzgraphiken

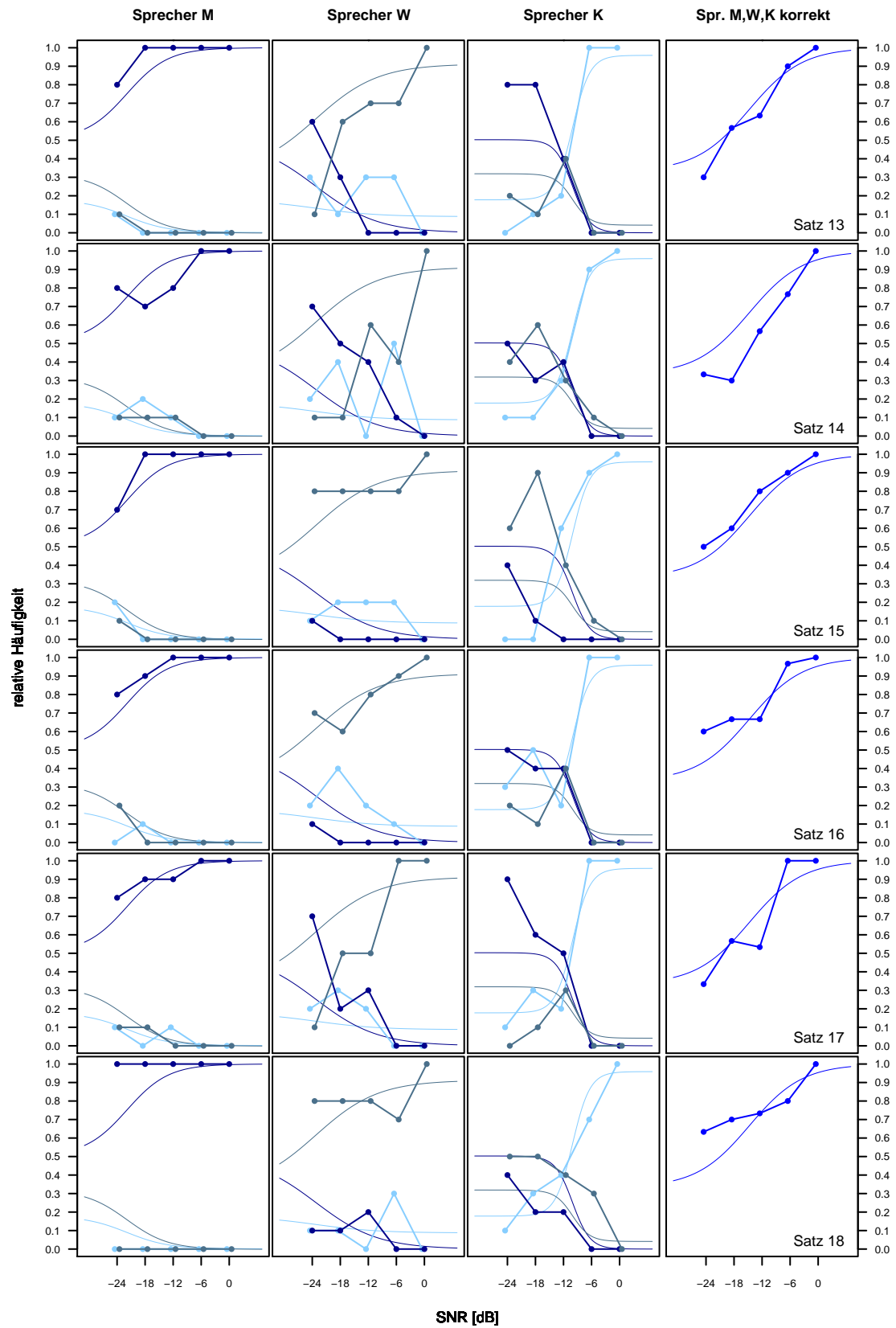


Abbildung C.9: Einzelsatzgraphiken der Testart Sprecher, Satz-Nr. 13–18

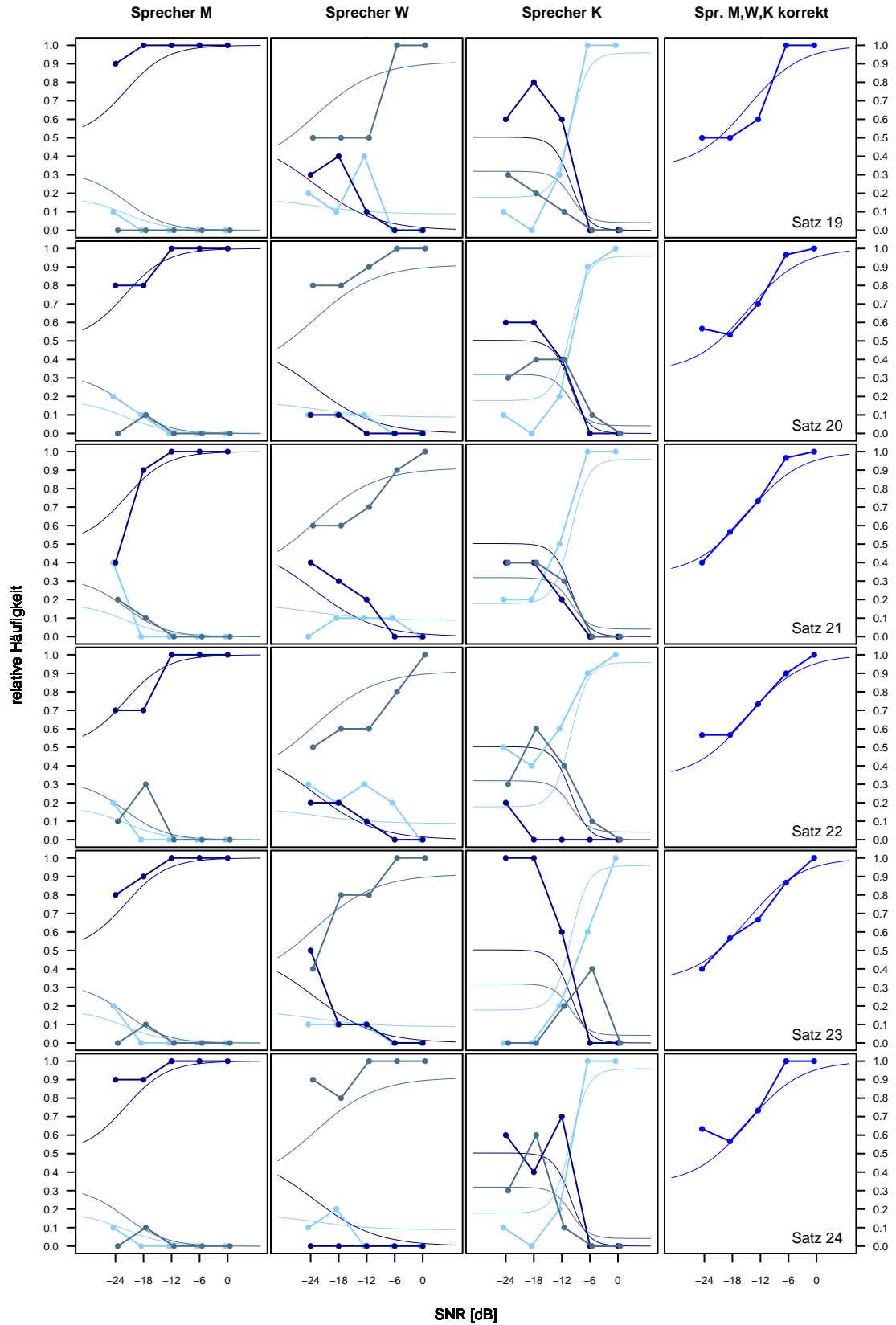


Abbildung C.10: Einzelsatzgraphiken der Testart Sprecher, Satz-Nr. 19–24

### C Einzelsatzgraphiken

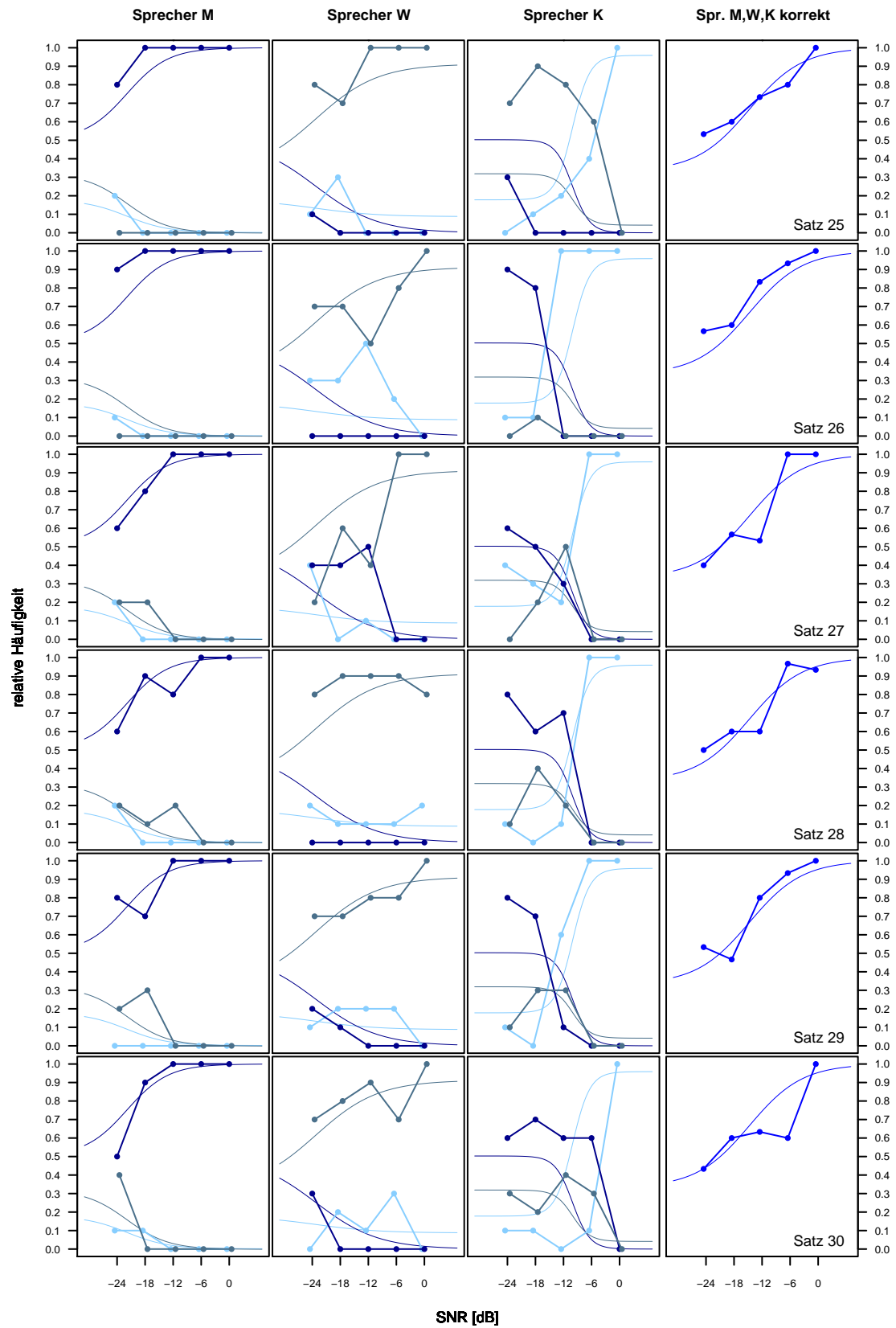


Abbildung C.11: Einzelsatzgraphiken der Testart Sprecher, Satz-Nr. 25–30

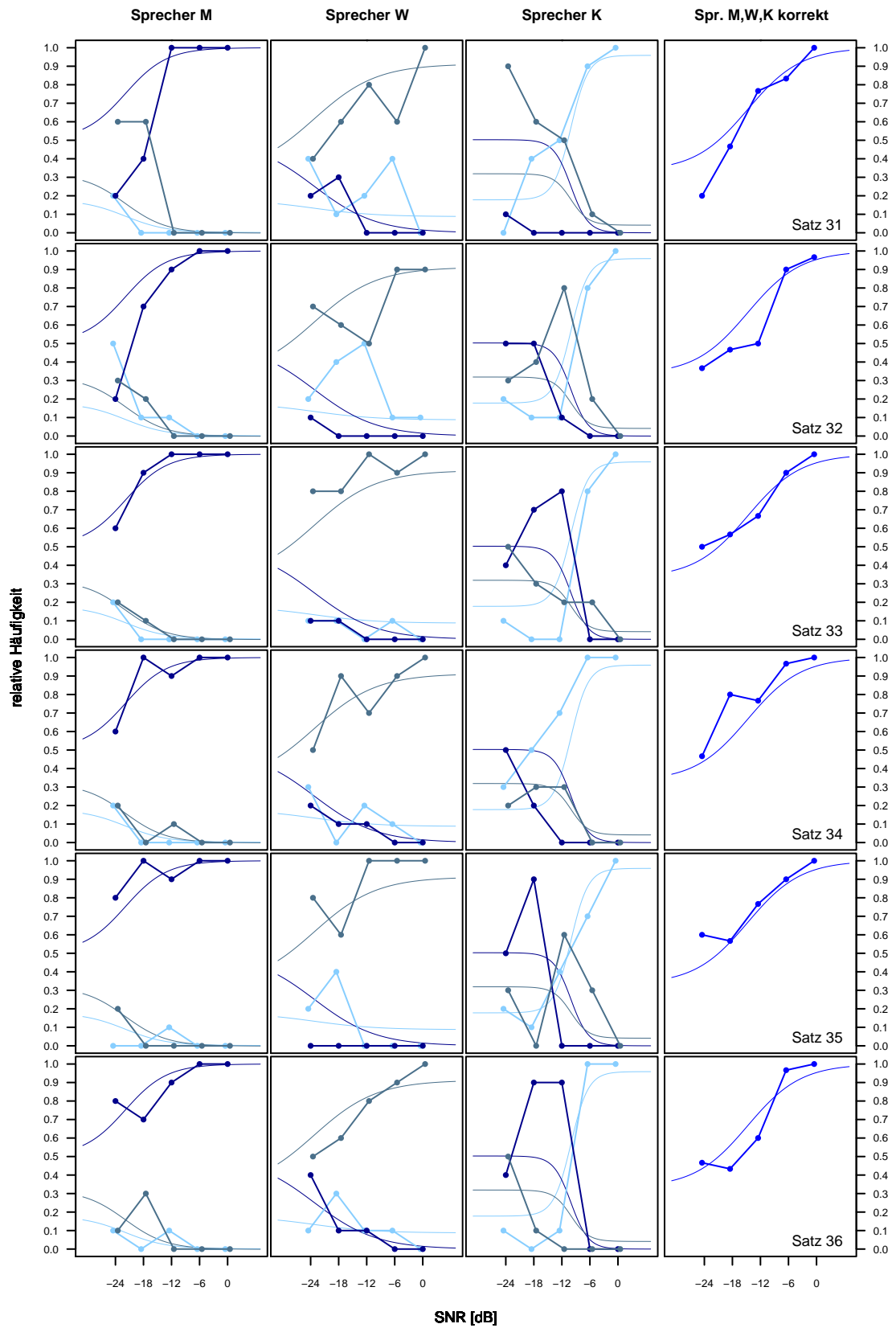


Abbildung C.12: Einzelsatzgraphiken der Testart Sprecher, Satz-Nr. 31–36





## D Satzvorschläge

Nach Niemeyer (1967), der schon 1963 Satztests für die Sprachaudiometrie entwickelte, ist darauf zu achten, dass die Sätze einfach, zwanglos und natürlich klingen. Sie sollen im Alltagsgespräch vorstellbar und „quasi selbstverständlich“ sein. Dies ist umso schwieriger, als dass diese Forderung in einem Prosodietest von jeweils drei Kategorien sowie Sprechern erfüllt werden muss. Die Sätze sollten in allen Kategorien und für alle Sprecher glaubwürdig sein. Um den Schwerhörigen nicht zu ermüden, sollten die Sätze und Satzgruppen kurz und handlich sein. Niemeyer (1967) schlägt vor, die Sätze aus vier bis sechs Wörtern zu erstellen und zehn Sätze zu acht Gruppen zusammenzufassen.

Des Weiteren sollten die Sätze untereinander keinen Sinnzusammenhang aufweisen. Vielmehr sollte der Inhalt stofflich weit gestreut sein, um eine Beeinflussung der Ergebnisse durch eine zufällige Vertrautheit oder Unbekanntheit des Satzinhalts möglichst zu vermeiden. Dabei sollten besonders „sprechertypische Begriffe“ vermieden werden (vgl. Abschnitt 4.4.2) und auf eine Eignung des Satzmaterials auch für Kinder geachtet werden. Außerdem sollte auf emotionsgefärbte Wörter möglichst verzichtet werden, um eine eventuelle Beeinflussung der Entscheidung für eine Satzkategorie auszuschließen (vgl. Abschnitt 4.4.1).

Nach Niemeyer (1967) müssen die Satzgruppen untereinander phonetisch ausbalanciert sein, zugleich aber auch ein Miniaturbild der Alltagssprache darstellen. Zur Annäherung an eine phonetische Ausbalanciertheit orientierte er sich an der prozentualen Lauthäufigkeitsverteilung nach Meier (in Niemeyer, 1967). Mögliche Auswirkungen verschiedener Laute auf das Testergebnis wurden ebenfalls speziell bei der Testart Sprecher der vorliegenden Arbeit beobachtet (s. Abschnitt 4.4.2). Aus diesem Grund wurde bei der Erstellung neuer Sätze versucht, auch dieses Kriterium mit einzubeziehen. Das erwies sich als teilweise schwierig.

## *D Satzvorschläge*

Im folgenden sind einige mögliche Sätze, die weitestgehend nach den genannten Kriterien erstellt wurden, aufgelistet:

- Tom hat keine Hausaufgaben auf
- Das Fahrrad ist neu
- Alle Kinder essen Süßes
- Das Spiel war teuer
- Das Bad ist schmutzig
- Uli kocht nur selten
- Karolin will tanzen lernen
- Die Uhr geht nach
- Anne isst keinen Blumenkohl
- Man sollte täglich Sport treiben
- Da ist die Zahnbürste
- Der Wecker klingelt morgens
- Man muss die Pflanzen gießen
- Alte Leute brauchen Pflege

## Literaturverzeichnis

- Adriaens Leon M.H. Ein Modell deutscher Intonation. In Inozuka (2003), Kapitel 2.4, Seite 44.
- Artemov V. A. Eksperimental naja fonetika. In Stock (1980), Kapitel 4.2.1, Seite 54.
- Bauer Monika. Zum Erkennen von Stimmqualitäten und prosodischen Mustern bei cochleaimplantierten Patienten. Julius-Maximilians-Universität Würzburg, Masterarbeit, 2005.
- Bilger R., Rzczekowski C. und Nuetzel J. et al. Evaluation of a test of speech perception in noise (SPIN). *A journal of the American Speech-Language-Hearing-Association / Asha*, 21:732, 1979.
- Boenninghaus und Lenarz . *Hals-Nasen-Ohren-Heilkunde für Studierende der Medizin*. Springer Verlag, 12. Ausgabe, 2005.
- Boothroyd A. Auditory perception of speech contrasts by subjects with sensorineural hearing loss. *J Speech Hear Res*, 27(1):134–144, Mar 1984.
- Boothroyd A. Perception of speech pattern contrasts via cochlea implants and limited hearing. *Ann Otol Rhinol Laryngol* 96, 128:58–62, 1987.
- Brill S. M., Gstöttner W., Helms J., Ilberg C. von, Baumgartner W., Müller J. und Kiefer J. Optimization of channel number and stimulation rate for the fast continuous interleaved sampling strategy in the COMBI 40+. *Am J Otol*, 18(6 Suppl):S104–S106, Nov 1997.
- Darwin C. J., Brungart D. S. und Simpson B. D. Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. *J Acoust Soc Am*, 114:2913–2922, 2003.
- Delattre P., Poenack E. und Olsen C. Some Characteristics of German Intonation for the Expression of Continuation and Finality. In Inozuka (2003), Kapitel 1.3, Seiten 14, 15.

- Dorman M. F., Loizou P. C. und Fitzke J. The identification of speech in noise by cochlear implant patients and normal-hearing listeners using 6-channel signal processors. *Ear Hear*, 19(6):481–484, Dec 1998a.
- Dorman M. F., Loizou P. C., Fitzke J. und Tu Z. The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6-20 channels. *J Acoust Soc Am*, 104(6):3583–3585, Dec 1998b.
- Essen O. von. Grundzüge der hochdeutschen Satzintonation. In Inozuka (2003), Kapitel 1.2, Seiten 10–14.
- Fery C. German Intonational Patterns. In Inozuka (2003), Kapitel 1.11, Seiten 27, 28.
- Fishman K. E., Shannon R. V. und Slattery W. H. Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. *J Speech Lang Hear Res*, 40(5):1201–1215, Oct 1997.
- Fu Q. J. und Shannon R. V. Phoneme recognition by cochlear implant users as a function of signal-to-noise ratio and nonlinear amplitude mapping. *J Acoust Soc Am*, 106(2):L18–L23, Aug 1999.
- Fu Q. J., Shannon R. V. und Wang X. Effects of noise and spectral resolution on vowel and consonant recognition: acoustic and electric hearing. *J Acoust Soc Am*, 104(6):3586–3596, Dec 1998.
- Fu Qian-Jie, Chinchilla Sherol und Galvin John J. The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users. *J Assoc Res Otolaryngol*, 5(3):253–260, Sep 2004. doi: 10.1007/s10162-004-4046-1. URL <http://dx.doi.org/10.1007/s10162-004-4046-1>.
- Fu Qian-Jie, Chinchilla Sherol, Nogaki Geraldine und Galvin John J. Voice gender identification by cochlear implant users: the role of spectral and temporal resolution. *J Acoust Soc Am*, 118(3 Pt 1):1711–1718, Sep 2005.
- Garnham Carolyn, O'Driscoll Martin, Ramsden Richard und Saeed Shakeel. Speech understanding in noise with a Med-El COMBI 40+ cochlear implant using reduced channel sets. *Ear Hear*, 23(6):540–552, Dec 2002. doi: 10.1097/01.AUD.0000042224.42442.A5. URL <http://dx.doi.org/10.1097/01.AUD.0000042224.42442.A5>.

- Hart Johan ´t, Collier René und Cohen Antonie. A Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody. In Inozuka (2003), Kapitel 2.4, Seite 44.
- Hellbrück Jürgen, Oguey Marie-Christin und Seiler Christian. Sind geschlechtsspezifische Unterschiede in der Lautstärkenempfindung ein Artefakt der Gehörganggröße? *Zeitschrift für experimentelle und angewandte Psychologie*, XXXI(3):439–446, 1984.
- Inozuka Emiko. *Grundzüge der Intonation. Definition und Methodologie in deutschen Intonationsmodellen*. Gunter Narr Verlag Tübingen, 2003.
- International Telecommunication Union . ITU-T Recommendation G.227; Conventional Telephone Signal. Blue Book, Fascicle III.2, 11 1988.
- Jones D. An Outline of English Phonetics. In Inozuka (2003), Kapitel 1.1, Seiten 9, 10.
- Kalikow D. N., Stevens K. N. und Elliott L. L. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *J Acoust Soc Am*, 61(5):1337–1351, May 1977.
- Kiss I. und Ennis T. Age-related decline in perception of prosodic affect. *Appl Neuropsychol*, 8(4):251–254, 2001.
- Kong Ying-Yee, Cruz Rachel, Jones J. Ackland und Zeng Fan-Gang. Music perception with temporal cues in acoustic and electric hearing. *Ear Hear*, 25(2):173–185, Apr 2004.
- Kraljic Tanya und Brennan Susan E. Prosodic disambiguation of syntactic structure: for the speaker or for the addressee? *Cognit Psychol*, 50(2):194–231, Mar 2005. doi: 10.1016/j.cogpsych.2004.08.002. URL <http://dx.doi.org/10.1016/j.cogpsych.2004.08.002>.
- Leder S. B., Spitzer J. B., Milner P., Flevaris-Phillips C., Richardson F. und Kirchner J. C. Reacquisition of contrastive stress in an adventitiously deaf speaker using a single-channel cochlear implant. *J Acoust Soc Am*, 79(6):1967–1974, Jun 1986.
- Lehiste I. Rhythmic units and syntactic units in production and perception. *J Acoust Soc Am*, 54(5):1228–1234, Nov 1973.

Lehiste I. und Meltzer D. Vowel and speaker identification in natural and synthetic speech. *Lang Speech*, 16(4):356–364, 1973.

Lehiste I. und Peterson . Some basic considerations in the analysis of intonation. In *Inozuka (2003)*, Kapitel 2.4, Seite 43.

Liu Chang und Kewley-Port Diane. Factors affecting vowel formant discrimination by hearing-impaired listeners. *J Acoust Soc Am*, 122(5):2855–2864, Nov 2007. doi: 10.1121/1.2781580. URL <http://dx.doi.org/10.1121/1.2781580>.

Machin D. und Campbell M. J. *Design of Studies for Medical Research*. John Wiley & Sons, Ltd, 2005.

Magne Cyrille, Schön Daniele und Besson Mireille. Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *J Cogn Neurosci*, 18(2):199–211, Feb 2006. doi: 10.1162/089892906775783660. URL <http://dx.doi.org/10.1162/089892906775783660>.

Meister H., Pyschny V., Landwehr M., Wagner P., Walger M. und Wedel H. von. [conception and realisation of a prosody test battery]. *HNO*, Jun 2007a. doi: 10.1007/s00106-007-1581-1. URL <http://dx.doi.org/10.1007/s00106-007-1581-1>.

Meister H., Tepeli D., Wagner P., Hess W., Walger M., Wedel H. von und Lang-Roth R. [experiments on prosody perception with cochlear implants]. *HNO*, 55(4):264–270, Apr 2007b. doi: 10.1007/s00106-006-1452-1. URL <http://dx.doi.org/10.1007/s00106-006-1452-1>.

Mitterbacher André. *Neue Signalverarbeitungs- und Stimulationsstrategien für Cochlea Implantate*. Dissertation, Technische Universität Wien, 2004.

Möbius B. Ein quantitatives Modell der deutschen Intonation. Analyse und Synthese von Grundfrequenzverläufen. In *Inozuka (2003)*, Kapitel 1.12, 2.4, Seiten 28–30, 43, 44.

Nie Kaibao, Barco Amy und Zeng Fan-Gang. Spectral and temporal cues in cochlear implant speech perception. *Ear Hear*, 27(2):208–217, Apr 2006. doi: 10.1097/01.aud.0000202312.31837.25. URL <http://dx.doi.org/10.1097/01.aud.0000202312.31837.25>.

- Niemeyer W. Sprachaudiometrie mit Sätzen. I. Grundlagen und Testmaterial einer Diagnostik des Gesamtsprachverständnisses. *HNO*, 15(11):335–343, Nov 1967.
- Owens E., Kessler D. K., Telleen C. C. und Schubert E. D. The Minimal Auditory Capabilities (MAC) Battery. *Hearing Aid Journal*, 34:9, 32,34, 1982.
- Peppé S., Maxim J. und Wells B. Prosodic variation in southern British English. *Lang Speech*, 43(Pt 3):309–334, 2000.
- Peppé Sue und McCann Joanne. Assessing intonation and prosody in children with atypical language development: the PEPS-C test and the revised version. *Clin Linguist Phon*, 17(4-5):345–354, 2003.
- Peterson G. H. und Barney H. L. Control methods used in a study of the vowels. *J Acoust Soc Am*, 24:175–184, 1952.
- R Development Core Team . *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2003. ISBN 3-900051-00-3.
- Richardson L. M., Busby P. A., Blamey P. J. und Clark G. M. Studies of prosody perception by cochlear implant patients. *Audiology*, 37(4):231–245, 1998.
- Ruhe C. Kommunikationsräume - auch für Hörgeschädigte! Technischer Bericht, 1999.
- Schatzer R., Kals M. und Zierhofer C. Encoding fine time structure with CSSS: concept and first results. *Wien Med Wochenschr.*, 156 (Suppl 119):93–4, 2006.
- Schumacher M. und Schulgen G. *Methodik klinischer Studien*. Springer-Verlag Berlin Heidelberg, 2007.
- Shatzman Keren B und McQueen James M. Prosodic knowledge affects the recognition of newly acquired words. *Psychol Sci*, 17(5):372–377, May 2006. doi: 10.1111/j.1467-9280.2006.01714.x. URL <http://dx.doi.org/10.1111/j.1467-9280.2006.01714.x>.
- Stelmachowicz P. G., Lewis D. E., Kelly W. J. und Jesteadt W. Speech perception in low-pass filtered noise for normal and hearing-impaired listeners. *J Speech Hear Res*, 33(2):290–297, Jun 1990.

- Stock E. *Untersuchungen zu Form, Bedeutung und Funktion der Intonation im Deutschen*. Akademie-Verlag Berlin, 1980.
- Waltzman S. und Hochberg I. Perception of speech pattern contrasts using a multichannel cochlear implant. *Ear Hear*, 11(1):50–55, Feb 1990.
- Westra GmbH . *Gerätebeschreibung für das Sprachaudiometer CAD 03/1*.
- Wilson B. S., Finley C. C., Lawson D. T., Wolford R. D., Eddington D. K. und Rabowitz W. M. Better speech recognition with cochlear implants. *Nature*, 352(6332): 236–238, Jul 1991. doi: 10.1038/352236a0. URL <http://dx.doi.org/10.1038/352236a0>.
- Wilson B. S., Finley C. C., Lawson D. T., Wolford R. D. und Zerbi M. Design and evaluation of a continuous interleaved sampling (CIS) processing strategy for multichannel cochlear implants. *J Rehabil Res Dev*, 30(1):110–116, 1993.
- Wilson Blake S und Dorman Michael F. Cochlear implants: a remarkable past and a brilliant future. *Hear Res*, 242(1-2):3–21, Aug 2008. doi: 10.1016/j.heares.2008.06.005. URL <http://dx.doi.org/10.1016/j.heares.2008.06.005>.
- Wingfield A., Lindfield K. C. und Goodglass H. Effects of age and hearing sensitivity on the use of prosodic information in spoken word recognition. *J Speech Lang Hear Res*, 43(4):915–925, Aug 2000.
- Wink K. *Wie liest und bewertet man eine klinische Studie?* Schattauer GmbH, Stuttgart, 2006.
- Zierhofer C. Electrical nerve stimulation based on channel specific sampling sequences, U.S. Patent 7,209,789, April 2007.



## Danksagung

Herrn Prof. Dr. med. Rudolf Hagen, Direktor der Klinik und Poliklinik für Hals-, Nasen- und Ohrenkrankheiten, plastische und ästhetische Operationen der Universität Würzburg und Herrn Prof. Dr. med. Joachim Müller, Oberarzt an der HNO-Klinik der Universität Würzburg bin ich für die Überlassung des Themas dieser Arbeit zu besonderem Dank verpflichtet.

Frau Prof. Dr. rer. nat. Kathleen Wermke danke ich für die Übernahme des Korreferats.

Herrn Dipl.-Ing. Stefan Brill danke ich für die Vergabe des Dissertationsthemas sowie die sehr gute Betreuung und Unterstützung in allen Belangen und während aller Phasen der Arbeit. Nur mit seiner Hilfe war es mir möglich, die technischen Schwierigkeiten bei der Vorbereitung des Tests zu bewältigen und die Ergebnisse statistisch auszuwerten und zu interpretieren.

Des weiteren möchte ich allen Mitarbeitern der CI-Abteilung sowie den Mitarbeitern der Audiometrieabteilung für die Unterstützung danken.

Zu besonderem Dank bin ich allen Probanden verpflichtet, die viel Zeit investierten und viel Geduld während der Versuche bewiesen haben.