

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

Direktor: Prof. Dr. med. Dr. h. c. Rudolf Hagen

Evaluation
des erweiterten HSM-Satztestmaterials
und Entwicklung einer Methode
für die Zusammenstellung
gleichwertiger Listen

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

Nicolai Lenz

aus Lörrach

Würzburg, November 2011

Referent: Prof. Dr. med. Dr. h. c. Rudolf Hagen

Korreferent: PD Dr.-Ing. Mario Cebulla

Dekan: Prof. Dr. med. Matthias Frosch

Tag der mündlichen Prüfung: 7. November 2012

Der Promovend ist Arzt

Meinen Eltern

Inhaltsverzeichnis

1	Einleitung	1
1.1	Hintergrund	1
1.2	Testverfahren	2
1.2.1	Einführung	2
1.2.2	Historie der Testverfahren und deren Kritikpunkte	2
1.2.3	Der HSM-Satztest	5
1.3	Aufgabe der Arbeit	6
2	Versuchsplanung und Methodik	7
2.1	Sprachmaterial	7
2.2	Versuchspersonen	10
2.3	Versuchsdurchführung	11
2.3.1	Versuchsaufbau	11
2.3.2	Satzdarbietung	12
2.3.3	Computerprogramm	13
2.4	Methode der Auswertung für die Listenzusammenstellung	15
3	Ergebnisse	18
3.1	Auswertungsdaten	18
3.2	Häufigkeitsverteilung der Sätze nach Verständniswert	18
3.3	Einflussfaktoren auf das Satzverständnis	22
3.3.1	Abhängigkeit von der Satzlänge	23
3.3.2	Abhängigkeit von der Satzkategorie	23
3.3.3	Abhängigkeit von der Zeitposition	27
3.3.4	Abhängigkeit von der Formelhaftigkeit	27
3.3.5	Abhängigkeit vom Geschlecht der Probanden	28
3.3.6	Abhängigkeit vom Inhalt der Satzblöcke	29
3.4	Listenzusammenstellung	31

4 Diskussion	33
4.1 Überblick	33
4.2 Probandenklientel	33
4.3 Satzgruppen in anderen Arbeiten	34
4.4 Satzverständlichkeit	34
4.4.1 Allgemein	34
4.4.2 Einfluss der Satzlänge	36
4.4.3 Einfluss der Satzkategorie	37
4.4.4 Einfluss der Zeitposition	38
4.4.5 Einfluss der Formelhaftigkeit	38
4.4.6 Einfluss des Inhalts der Satzblöcke	40
4.5 Satzlautstärke und deren Einfluss	40
4.5.1 Beobachtung	40
4.5.2 Einfluss durch Aufsprache	41
4.5.3 Einfluss durch Lautstärkenangleichung	44
4.5.4 Fazit	44
4.6 Listenzusammenstellung	45
5 Zusammenfassung	46
Literaturverzeichnis	47
Danksagung	
Lebenslauf	

1 Einleitung

1.1 Hintergrund

Das Hörorgan ist eines der wichtigsten Sinnesorgane des Menschen. Es befähigt uns unsere Umwelt wahrzunehmen. Wir hören Töne, Klänge, Geräusche und auch ein komplexes Frequenzgemisch wie Sprache.

Diese ermöglicht es uns, uns mitzuteilen und außerdem durch Unterschiede in der Sprachdynamik, der Sprachmelodie, der Klangfärbung und auch des Sprachrhythmus zusätzliche Informationen zu geben (Niemeyer, 1967; Niemeyer und Beckmann, 1962). Erst durch das Hörvermögen haben wir die Möglichkeit das Sprechen zu erlernen. Somit ist das Hören Grundlage sowohl für die Informationsaufnahme, als auch für die verbale Informationsabgabe.

Auf dieser Kommunikationsfähigkeit basiert wiederum die soziale Integration in die Gesellschaft. Sei es im familiären Umfeld, sei es in der Arbeitswelt, im direkten Kontakt oder am Telefon, überall ist Informationsaustausch wichtig.

Um so schwieriger stellt sich die Situation für all die Menschen dar, die durch eine Ertaubung nicht die Möglichkeit haben an der auf Sprache basierenden zwischenmenschlichen Kommunikation teilzunehmen. Dazu zählen diejenigen, die durch pränatale Erkrankungen oder Entwicklungsstörungen taub zur Welt kommen, aber auch diejenigen, die auf Grund einer Erkrankung im Laufe des Lebens ertauben. Bei letzteren ist zumindest die Sprechfähigkeit vorhanden.

Lange Zeit war es nicht möglich diesen Menschen zu helfen. Sie mussten lernen mit ihrer Behinderung zu leben und die Welt der Gehörlosigkeit zu akzeptieren. Dies änderte sich als im Jahre 1957 das erste Cochlea-Implantat (CI) eingesetzt wurde. Seitdem ist die Entwicklung immer weiter vorangeschritten und seit Mitte der 80er Jahre hat die CI-Versorgung klinische Routine erlangt. Bis heute sind ertaubte Menschen weltweit mit insgesamt über 200000 Cochlea-Implantaten versorgt worden. Liegt eine beidseitige cochleäre Ertaubung vor, werden heutzutage beide Seiten mit einem CI versorgt. Es haben sich in verschiedenen Studien eindeutige Vorteile vor allem im Richtungshören und beim Hören mit Hintergrundgeräuschen für diese Patienten gezeigt (Dunn et al., 2008;

Laszig et al., 2004; Müller et al., 2002). Die Betroffenen erreichen damit eine Hörleistung von bis zu 80 Prozent eines Normalhörenden. Sie können Musik hören, telefonieren und Kinder können Schulen für Normalhörende besuchen (Helms und Müller, 1999).

Die Patientenzahl nimmt immer mehr zu und gerade prälingual Ertaubte werden in einem immer jüngeren Alter operiert. Dies bedeutet, dass die CI-Betreuung über einen zunehmend längeren Zeitraum durchgeführt werden muss. In der Regel dauert es drei Jahre bis die Anpassung bei einem prälingual ertaubten Kind abgeschlossen wird und in der Folgezeit ist eine halbjährliche Überprüfung empfohlen. Für die Anpassung wie für die Kontrollen gibt es verschiedene Testverfahren.

1.2 Testverfahren

1.2.1 Einführung

Zum einen gibt es tonaudiometrische Verfahren zur Ermittlung der Hörschwelle. Diese Testergebnisse treffen eine Aussage über die akustische Wahrnehmungsschwelle des Patienten; allerdings ist damit nicht das Verstehen von Worten und Sätzen garantiert. Hierfür gibt es eine ganze Reihe von Sprachverständnistests.

Je sensitiver diese Testverfahren sind, desto besser lassen sich die CIs einstellen. Dadurch wird die Situation eines CI-Patienten idealerweise immer mehr der eines Normalhörenden angenähert. Um dies zu gewährleisten, unterliegen die Sprachverständlichkeitstests einem fortlaufenden Entwicklungsprozess.

1.2.2 Historie der Testverfahren und deren Kritikpunkte

Es wurden im Laufe des letzten Jahrhunderts eine ganze Reihe von Sprachtests entwickelt. Mit zunehmender Erfahrung wurden manche von ihnen wieder verworfen, andere weiterentwickelt. Der folgende Abschnitt soll einen kurzen Überblick über die Entwicklung derartiger Sprachtests geben und auf die jeweiligen Erkenntnisse und Kritikpunkte, die sich daraus ergaben, eingehen.

Der Freiburger Test, der von Hahlbrock (1957) entwickelt wurde, war das erste deutsche sprachaudiometrische Testverfahren. Hierbei wurden dem Probanden Zahlen und einsilbige Worte dargeboten. Mit diesem Test wurden zwei verschiedene Eigenschaften gemessen: die Sprachverständnisschwelle und das

Diskriminationsvermögen. Mit den Zahlen war es möglich die Sprachverständnisschwelle zu bestimmen. Da Zahlen jedoch eine hohe Redundanz haben, ist es schwierig eine Aussage über das Diskriminationsvermögen des Probanden zu treffen. Hierfür eignet sich der Einsilber-Test besser.

In der Folgezeit wurde dieser Test in mehrfacher Hinsicht kritisiert. Bangert (1980) zeigte, dass die einzelnen Testmaterial-Gruppen untereinander sehr inhomogen waren, so dass sich die Werte für die Diskriminationsleistung einer einzelnen Versuchsperson von Testmaterial-Gruppe zu Testmaterial-Gruppe sehr stark unterschieden. Auch die Reproduzierbarkeit der Ergebnisse in einer Gruppe sei nicht zuverlässig möglich. Dieser Kritik schloss sich auch von Wedel (1986) an.

Niemeyer und Beckmann (1962) wiesen darauf hin, dass der Test nicht auf den Alltag übertragen werden könne. Zum einen würde ein Wort in einem Satz eingebettet zusätzliches Informationsmaterial beinhalten, anders als bei einem Einzelwort und zum anderen sind in Alltagssituationen Störgeräusche vorhanden, die in diesem Test fehlen. Doch genau der Gebrauch der Sprache im Alltag ist der entscheidende Faktor für den Patienten.

Auf diesen Erkenntnissen basierend entwickelte Niemeyer den Marburger Satztest (Niemeyer und Beckmann, 1962). Er enthielt 8 Gruppen mit je 10 Sätzen, die verrauscht und unverrauscht dargeboten werden konnten. Mit diesem Test zeigte er, dass bei 50prozentiger Einsilberdiskriminationsfähigkeit bei Normalhörenden immer noch ein 80-90prozentiges Satzverständnis vorlag. Dies gelte aber nicht für Schwerhörige. Anhand von konkreten Beispielen legte er dar, dass bei Schwerhörigen, bei denen jeweils ähnliche Zahlen- und Einsilberverständlichkeiten vorlagen, sich das Satzverständnis deutlich unterschied. Auch Schultz-Coulon (1974) befasste sich mit dieser Thematik und stellte fest, dass gerade für Patienten mit Hochtonverlust keine lineare Beziehung bestand. Bei dieser Klientel wirkten sich die Störgeräusche extrem auf die Sprachverständlichkeit aus.

In einer weiteren Arbeit befasste sich Niemeyer (1967) genauer mit den Unterschieden von Schallempfindungs- und Schalleitungsschwerhörigkeit in Sprachtests. Als Ergebnis hielt er fest, dass sich für Normalhörende – mit Einschränkung auch für Schalleitungsgestörte – eine Korrelation zwischen Einsilberverständnis und Satzverständnis zeigen lässt, jedoch nicht für Personen mit Schallempfindungsstörungen. Er betonte, dass der Satztest ein ergänzendes Mittel in der Sprachaudiometrie darstelle und nicht Zahlen- und Einsilbertest ablösen

könne. Zur orientierenden Diagnostik sind diese Mittel weiterhin probat und haben auch bis heute in der Klinik Relevanz.

In der Folgezeit wurde auch der Marburger Satztest kritisiert. Es wurde bemängelt, dass er zu wenig Sprachmaterial enthalte und daher nicht reproduzierbar sei (Schultz-Coulon, 1974). Außerdem sei heutzutage das verwendete Vokabular zu „altertümlich“ (Lehnhardt und Laszig, 2009).

1997 wurde der Göttinger Satztest von Kollmeier und Wesselkamp (1997) entwickelt. Er enthält 20 Listen zu je 10 fixen Sätzen zwischen 3 und 7 Wörtern. Mit diesem Satztest wird die Sprachverständlichkeitsschwelle (speech reception threshold; SRT) im Störgeräusch gesucht und mit dem Wert für Normalhörende verglichen.

Der Oldenburger Satztest, der von Wagener et al. (1999) entwickelt wurde, sucht ebenfalls den SRT bei einem Störgeräusch, er verwendet allerdings semantisch nicht vorhersagbare 5-Wort-Sätze. Jeder dieser Sätze besteht aus einem Namen, einem Verb, einem Zahlwort, einem Adjektiv und einem Objekt. Dabei werden diese Wortkombinationen willkürlich aus einem Gesamtwortschatz von 50 Wörtern zusammengestellt. Insgesamt beinhaltet der Test 40 Listen mit je 30 Sätzen. Durch die geringe Anzahl an Wörtern im Gesamtwortschatz ist eine starke Homogenität des Testes über alle Testlisten gegeben und durch die willkürliche Wortkombination in einem Satz ist die Vorhersagbarkeit und damit der Wiederholungseffekt gering. Jedoch handelt es sich bei zunehmender Anwendung um einen geschlossenen Test, da durch die zahlreichen Wiederholungen davon ausgegangen werden kann, dass alle 50 Wörter bekannt sind.

Sowohl der Göttinger Satztest als auch der Oldenburger Satztest haben einen festen Platz in der Nachsorge von CI-Patienten sowie in der Hörgeräteanpassung.

Auch der HSM-Satztest, der 1997 von Hochmair-Desoyer, Schultz und Moser (Hochmair-Desoyer et al., 1997) entwickelt wurde, basiert auf Listen mit fixen Sätzen zwischen 3 und 8 Wörtern Länge. Hier bilden 30 Listen mit je 20 Alltagssätzen die Grundlage. Die 20 Sätze pro Liste beinhalten jeweils 106 Wörter. Mit insgesamt 600 Sätzen ist die Gesamtwortmenge deutlich größer als bei den bisher vorgestellten Tests.

Vom Schwierigkeitsgrad her liegt dieser Satztest zwischen dem Göttinger und dem Oldenburger Satztest (Lehnhardt und Laszig, 2009). Bei diesem sprachaudiometrischen Testverfahren wird bei fest vorgeschriebenem Signal-Rauschabstand (signal to noise ratio; SNR) das Satzverstehen im Störschall getestet. Es

dient als Qualitätsmaß für den therapeutischen Erfolg von CI-Patienten.

Parallel zu den genannten Satztests wurden weitere Tests entwickelt, bei denen nur einzelne Teile des Satzes genauer betrachtet wurden. Tschopp und Ingold (1992) entwickelten eine deutsche Version des SPIN-Tests (Speech Perception in Noise; Sprachverstehen bei Hintergrundlärm). Bei dessen Anwendung wird der Proband gebeten das letzte Wort im Satz – ein einsilbiges Substantiv – wiederzugeben. Es wird im Gegensatz zum restlichen Satz verrauscht dargeboten. Es werden Substantive mit einer hohen und Substantive mit einer niedrigen Vorhersagewahrscheinlichkeit auf Grund des Kontextes unterschieden. Mit diesem Verfahren können Hinweise auf Störungen der zentral-auditiven Sprachverarbeitung gewonnen werden.

Außerdem wurden Testverfahren entwickelt, die Unterschiede auf Konsonanten- und Vokal-Ebene betrachten. Dazu gehört der 1982 von Sotschek (1982) entwickelte Reimtest. Hierbei handelt es sich um einen geschlossenen Test, d. h. dem Probanden sind die Antwortmöglichkeiten vorab bekannt. Ihm werden 6 Einsilber als Antwortmöglichkeit angeboten, die sich im Anlaut, Inlaut oder Auslaut unterscheiden. Dieser Test dient der Fehlhörigkeitsanalyse. Er trifft nicht nur eine Aussage über das prozentuale Verstehen, sondern man kann mit ihm auch eine Analyse der Phonem-Verwechslungen durchführen und damit Informationen über Störungen der akustischen Verarbeitung erhalten. Dieser Reimtest wurde in der Folgezeit von Kollmeier et al. (1992) zur klinischen Anwendung weiterentwickelt.

Demzufolge ist in den letzten Jahrzehnten die Entwicklung auf dem Gebiet der Sprachverständnistests stark vorangeschritten, auf die verschiedenen Kritikpunkte wurde mit Weiter- und Neuentwicklungen reagiert.

1.2.3 Der HSM-Satztest

Der HSM-Satztest gilt als ein sehr geeignetes Instrument für die Imitation der Alltagshörsituation bei CI-Patienten und wird klinisch häufig angewendet. Es wird aber auch zu diesem Test Kritik geäußert. So berichtet Müller-Deile (2009), dass beim HSM-Satztest ein sehr großer Wiedererkennungseffekt bestünde und empfiehlt bei der Testung mit derselben Satzliste mindestens ein Jahr Abstand zu lassen. Auch Mrowinski und Scholz (2006) berichten über unerwünschte Wiederholungseffekte. Da solche Tests auch zu wissenschaftlichen Zwecken, zum Beispiel zur Überprüfung der Lebensqualitätssteigerung durch CIs (Hirschfelder

et al., 2008) oder zur Objektivierung von Kurz- und Langzeitergebnissen mit CIs (Oh et al., 2003) genutzt werden und in diesem Rahmen Mehrfachtestungen von Patienten innerhalb kürzester Zeit vorkommen können, wird auch aus dieser Richtung Kritik laut (Brill, 2008). Weiterhin wird aus der klinischen Erfahrung mit dem HSM-Satztest berichtet, dass es möglicherweise Sätze gäbe, die einfacher seien als andere (Brill, 2008). Hierdurch wird die Objektivität des Testverfahrens in Frage gestellt. Objektive Testergebnisse sollten bei Fragen der Kostenübernahme durch Krankenkassen, aber auch im Rahmen der Berufsunfähigkeit bzw. Erwerbsminderung Voraussetzung sein.

All diese aufgeführten Gründe legen eine Erweiterung und Neustrukturierung des HSM-Satztests nahe. So beschäftigte sich Angermüller (unveröff. Diss.) mit der Neuaufsprache der ursprünglichen HSM-Sätze und der Aufsprache neu hinzukommender Sätze.

1.3 Aufgabe der Arbeit

Die hier vorliegende Arbeit befasst sich mit der Evaluation der von Angermüller (unveröff. Diss.) zugrunde gelegten erweiterten HSM-Sätze. Als erstes soll die Schwierigkeit eines jeden einzelnen Satzes untersucht und eingestuft werden. Dazu werden die Sätze in gleicher Lautstärke und mit einem Rauschen belegt hörgesunden Probanden dargeboten. Der jeweilige Proband gibt alle verstandenen Inhalte eines jeden Satzes wieder. Daraus ergibt sich, über alle Probanden gemittelt, für jeden Satz ein Schwierigkeitswert. Im Anschluss sollen Listen aus diesen Sätzen nach einem bestimmten Schema zusammengestellt werden, so dass diese Listen möglichst gleich schwierig und damit vergleichbar und austauschbar sind.

Hierbei sollen die von Wesselkamp et al. (1992) geforderten Kriterien für die Validität eines Satztests eingehalten werden.

Durch einen großen Umfang an Satzmaterial soll erreicht werden, dass in Zukunft Mehrfachmessungen mit dem HSM-Satztest über einen kürzeren Zeitraum durchgeführt werden können, ohne dass eine Verfälschung durch Wiederholungseffekte auftritt. Auch sollen durch die Neuzusammenstellung der Satzlisten mögliche Unterschiede im Schwierigkeitsgrad der bisherigen Satzzusammenstellungen verringert oder ganz vermieden werden. Somit ist es Ziel dieser Arbeit, einen erweiterten HSM-Satztest zu entwickeln, der eine höhere Validität besitzt.

2 Versuchsplanung und Methodik

2.1 Sprachmaterial

Grundlage für die Versuchsreihe bildeten die Sätze aus der Dissertation von Angermüller (unveröff. Diss.). Diese beinhaltete eine Aufsprache der originalen HSM-Sätze (Hochmair-Desoyer et al., 1997) und zusätzlich weitere neu hinzugefügte Sätze. Die Auswahlkriterien, die Aufspreekriterien und die Zusammenstellung der Sätze wurden in der genannten Dissertation (Angermüller (unveröff. Diss.)) festgelegt.

In der hier vorliegenden Arbeit wurden anfangs alle Sätze ein weiteres Mal zur Kontrolle abgehört. Inhaltlich sowie phonetisch fragwürdige Sätze wurden komplett aus dem Satzpool gestrichen oder gegebenenfalls durch eine andere Aufspracheversion ersetzt.

Der Satzpool beinhaltete schlussendlich 1388 Sätze mit einer Länge von 3 bis 8 Wörtern. Bei den Satzkategorien wurde zwischen Aussagesatz, Kommandosatz, Ausrufesatz und Fragesatz unterschieden. Aus Tabelle 2.1 auf S. 7 ist die Aufteilung der Sätze in Abhängigkeit der Wortanzahl zu den verschiedenen Satzkategorien zu entnehmen.

Diese 1388 Sätze wurden mit einem CCITT-Rauschen (International Telecommunication Union, 1988) belegt. Hierbei handelt es sich um ein international genormtes Rauschen, dessen Spektrum an die mittlere Häufigkeitsverteilung mehrerer Sprachen angepasst ist.

Tabelle 2.1: Aufteilung der Sätze nach Wortanzahl

	Aussagesatz	Fragesatz	Ausrufesatz	Kommando
132 Drei-Wortsätze	132	0	0	0
281 Vier-Wortsätze	0	131	68	83
450 Fünf-Wortsätze	155	136	81	77
264 Sechs-Wortsätze	131	133	0	0
130 Sieben-Wortsätze	130	0	0	0
131 Acht-Wortsätze	131	0	0	0

In Vorversuchen bei unterschiedlichen SNRs, die Angermüller (unveröff. Diss.) verwendete, zeigte sich, dass bei der Neuaufsprache ein 50prozentiges Verstehen bei einer SNR von -9,91 dB in der Probandengruppe erreicht wurde. Somit wurde ein SRT von -10 dB für die Versuchsdurchführung zu Grunde gelegt. Das bedeutete, dass zu der Lautstärke des Satzes noch ein Rauschen hinzugefügt wurde, so dass eine SNR von -10 dB resultierte.

Folgende Erläuterungen werden auch graphisch in Abbildung 2.1 auf S. 9 dargestellt. Am Anfang und am Ende jedes Satzes (a) wurden sinusförmige Rampen in der Länge von 50 ms eingeführt, um ein Knackgeräusch beim Übergang von der Stille zur Sprache beziehungsweise umgekehrt zu verhindern (b). Die vordere Rampe stellt eine von 0 nach 1 aufsteigende Sinuskurve dar, die hintere Rampe am Satzende analog dazu eine von 1 nach 0 absteigende Sinuskurve. Die Spanne von 0 bis 1 beinhaltet das Erreichen der wirklichen Lautstärke des Satzes. Die Spanne von 1 nach 0 demonstriert den Abfall von der Satzlautstärke bis zur absoluten Stille.

Für das Rauschen wurde gleichermaßen verfahren und am Satzanfang wie am Satzende ein über 300 ms zunehmendes beziehungsweise abnehmendes Rauschen in Sinusform hinzugefügt (e). Hierdurch wurden die Probanden weder am Satzanfang noch am Satzende mit einem plötzlichen Start bzw. Abbruch des Rauschens konfrontiert (f). Beide Signale wurden durch gewichtete Addition gemischt und sobald die maximale Lautstärke des Rauschens erreicht war, wurde mit der 50 ms langen sinusförmigen Rampe beginnend der eigentliche Satz dargeboten, so dass nun Satz und Rauschen gleichzeitig wahrgenommen wurden (g).

Des Weiteren wurden die Sätze auf dieselbe Lautstärke gebracht. Dazu wurde von Dipl.-Ing. Brill ein Matlab-Skript verfasst, was ähnlich wie ein Schallpegelmessgerät die verschiedenen Schalldruckpegel (sound pressure level; SPL; A-bewertet, „fast“) im Satzverlauf misst. Nach dem Messen der Schalldruckpegel aller Satzverläufe wurde der maximal mögliche Pegel – gültig für alle Sätze – berechnet. Der limitierende Faktor dabei war das Clipping. Dies galt es zu vermeiden, damit die Sätze im codierbaren Bereich lagen. Nach Ermittlung dieses Maximalpegels wurden alle Sätze durch Multiplikation auf die ermittelte Lautstärke berechnet und ein Rauschen aufsummiert, so dass eine SNR von -10 dB resultierte. Auf das Verfahren der Pegelberechnung wird im Ergebniskapitel (s. Abschnitt 4.5.1 auf S. 40) noch einmal genauer eingegangen.

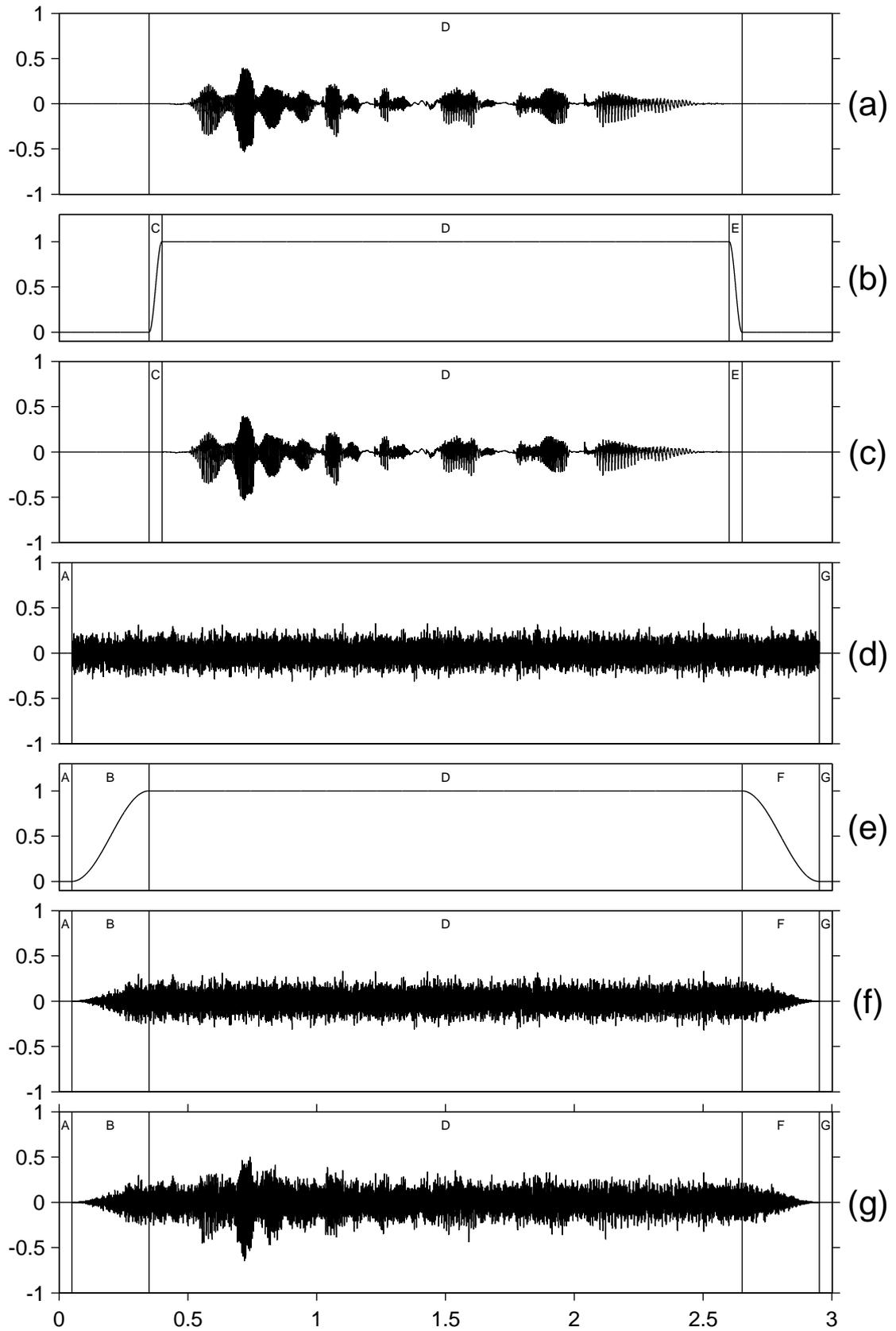


Abbildung 2.1: Erzeugung eines Satzes mit Störgeräusch

2.2 Versuchspersonen

Um das Kriterium einer möglichst homogenen Versuchspersonengruppe zu erfüllen, wurden die Probanden nach folgenden Vorgaben ausgesucht:

- Alter zwischen 20 und 30 Jahren
- Student/in
- deutsch als Muttersprache
- hörgesund (Selbsteinschätzung)
- unauffälliger Audiometriebefund

Hierzu wurde im Vorfeld der eigentlichen Untersuchung mit jedem Probanden eine wie im Klinikalltag übliche Audiometrie durchgeführt. Diese fand in der Hörkabine B der HNO-Klinik Würzburg statt, in der rechten der beiden Kammern. In Abbildung 2.2 auf S. 11 wurden die Audiometriebefunde aller Probanden übereinander gelegt. Hieraus war abzulesen, dass eine maximale Abweichung von 20 dB von der Normalhörenden-Kurve im Frequenzbereich von 250 – 6000 Hertz akzeptiert wurde. Somit ist der ISO-Norm 7029 (International Organization for Standardization, 2000), die sich mit der Normalhörigkeit in Abhängigkeit des Alters befasst, entsprochen. Ein Proband fiel durch besonders schlechte Ergebnisse auf, was sich auch in seinem Audiogramm zeigte. Seine Tonschwelle wich um bis zu 30 dB von der Referenzlinie ab. Somit entsprach der Proband nicht den vorgegebenen Kriterien und wurde in der weiteren Auswertung nicht berücksichtigt. Eine andere Testperson ersetzte diesen Probanden.

Die Anzahl der weiblichen und männlichen Probanden war nicht explizit vorgeschrieben. Es zeigte sich nach Durchführung aller Versuche, dass ein Geschlechterverhältnis von 1:1 vorlag. Es ist darauf hinzuweisen, dass für Messungen im Freifeld bisher keine geschlechtsspezifischen Unterschiede in der Literatur beschrieben worden sind. Solche sind lediglich für Lautheitsempfindungsmessungen mit Kopfhörern bekannt (Hellbrück et al., 1984), was auf die Größenunterschied der Gehörgänge von Männern und Frauen zurückgeführt wird.

Die Reihenfolge der zu testenden Probanden und damit die Zuordnung zu dem jeweiligen Testmaterial wurde durch Randomisierung festgelegt.

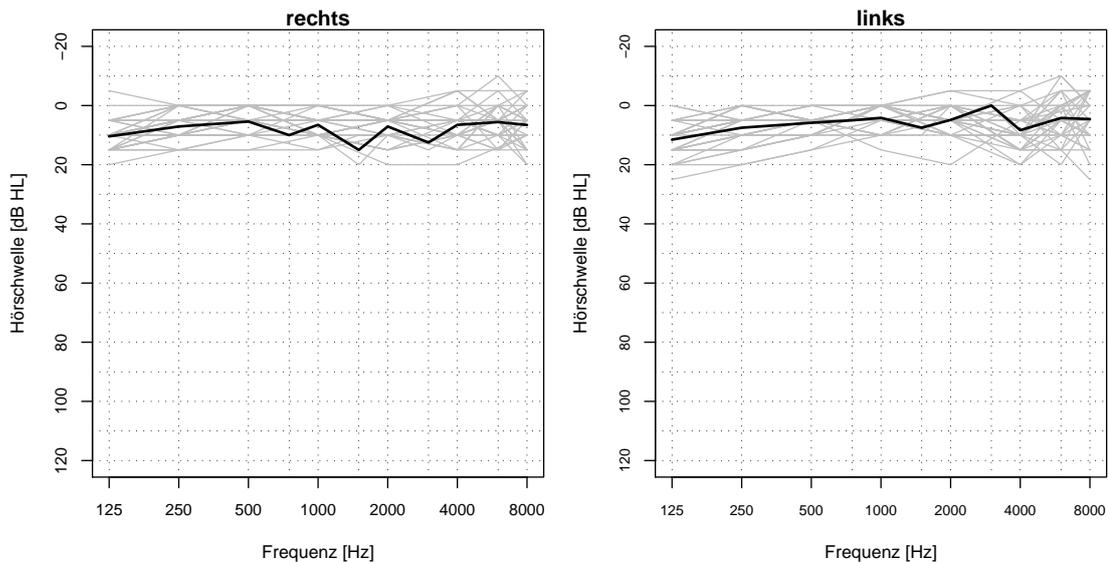


Abbildung 2.2: überlagert dargestellte Audiogramme (schwarze Linie: Mittelwert)

2.3 Versuchsdurchführung

Jeder Proband bekam eine feste Anzahl von Sätzen verrauscht dargeboten. Es war Aufgabe des Probanden das Gehörte so gut wie möglich wiederzugeben. Dabei wurde jedes richtige Wort als korrekt und jedes falsch oder nicht verstandene Wort als nicht korrekt gewertet.

2.3.1 Versuchsaufbau

Der Versuch selbst fand in der schallgedämpften Hörkabine C der HNO-Abteilung der Universität Würzburg statt. Der Proband nahm auf einem Stuhl Platz, der 1,20 m vom Lautsprecher entfernt war. Der Lautsprecher befand sich vor dem Probanden, so dass dieser der Schallquelle entgegensah. Mit diesen Voraussetzungen waren die von Lehnhardt und Laszig (2009) geforderten Bedingungen für die Freifeldmessung für Sprachaudiometrie erfüllt.

Der Untersucher war im selben Raum gegenüber dem Probanden hinter dem Lautsprecher an einem Audiometriemessplatz sitzend, um von dort den Versuch zu leiten.

Dort war der Audiometrie-PC, über den die Sounddateien im Wave-Format mit dem Winamp Media Player 2.71 abgespielt wurden. Des Weiteren war dort folgender Laptop zur Auswertung: ein hp Compaq nc 6320. Das Ausgangssignal wurde über die Audiometriebox WESTRA LAB 1001 zum Audiometrielautspre-

cher weitergeleitet. Ziel war es, dass das Signal beim Probanden mit 75 dB ankam – also 65 dB Sprachsignal plus zusätzlich einem Rauschen mit einer SNR von -10 dB. Um die richtige Einstellung für diesen Zielwert zu erreichen, wurde ein Schalldruckpegelmessgerät am Probandenplatz in Ohr-Höhe positioniert. So ermittelt, musste an der Audiometriebox ein Schalldruckpegel von nominell 83 dB bei standardisierter PC-Lautstärke eingestellt werden, um dem festgelegten Zielwert von 75 dB zu entsprechen. Außerdem musste die Einstellung „Freifeld“ gewählt werden.

2.3.2 Satzdarbietung

In der Versuchsreihe wurden den Probanden die verwaschenen Sätze dargeboten. Dabei waren die Sätze im Vorhinein schon mit einem Rauschen – wie oben beschrieben – versehen worden. Aus zeitlichen Gründen sowie aus Gründen der Konzentration und Belastung war es nicht möglich, jedem Probanden alle 1388 Sätze zur Beurteilung vorzuspielen. Die 1388 Sätze wurden in 12 Blöcke eingeteilt: 8 Blöcke zu je 116 Sätzen und 4 Blöcke zu je 115 Sätzen. Jedem Probanden wurden 4 Blöcke vorgespielt, zwischen denen jeweils eine kurze Pause gemacht werden konnte. Dies entsprach etwa dem Umfang eines Drittels aller Sätze.

Die Verteilung der Sätze auf die 12 Blöcke wurde willkürlich nach dem Zufallsprinzip durchgeführt. Bei den genannten Rahmenbedingungen wurde eine Anzahl von 36 Probanden festgelegt, so dass jeder Satz 12 mal von jeweils unterschiedlichen Probanden gehört wurde.

Wie oben erwähnt, musste ein Proband - Nummer 2 - ausgeschlossen und durch einen nachrückenden Probanden - Nummer 37 - ersetzt werden. Dieser bekam exakt dasselbe Testmaterial dargeboten wie der ursprüngliche Proband Nummer 2. Die Durchnummerierung der Probanden folgte den Positionen 1 bis 37. Für die Auswertung wurden nur die Ergebnisse der Probanden 1 und 3 bis 37 verwendet, so dass im Folgenden den Ergebnissen der Probanden 3 bis 37 die Nummern 2 bis 36 zugewiesen wurden.

Für die Blockdarbietung wurde ein permutierendes System angewendet. Die Blöcke wurden mit den Zahlen 1 bis 12 durchnummeriert. Der erste Proband bekam die Blöcke 1-4 dargeboten, der zweite die Blöcke 2-5. Analog dazu wurde für die folgenden Probanden fortgefahren: Für jeden neuen Probanden wurde die Blockreihenfolge um eins nach rechts verschoben. Sobald Block 12 erreicht war, wurde erneut mit Block 1 angefangen. Die Zuordnung der Blöcke auf die

Probanden ist als Tabelle in Abbildung 2.2 auf S. 14 dargestellt.

Anfänglich hörte jeder Proband einen Probekblock ab, um sich auf die Bedingungen einzustellen. Dieser Probekblock beinhaltete einen Ausschnitt von 20 Sätzen aus einem nicht für den Probanden vorgesehenen Block.

Durch die beschriebene Anordnung der Blöcke wurde garantiert, dass derselbe Block für unterschiedliche Probanden unterschiedliche Positionen in der Blockreihenfolge einnahm. Durch diese variierende Stellung eines Blockes sollten mögliche Unterschiede beispielsweise durch Konzentrationsgefälle und Gewöhnungsaspekte vermieden werden. Zur weiteren Vermeidung von Effekten, die in einem Zusammenhang mit der Positionierung stehen, wurden für alle Probanden, die denselben Block dargeboten bekamen, die Sätze jeweils zufällig innerhalb eines Blockes neuverteilt.

2.3.3 Computerprogramm

Die zur Versuchsdurchführung benötigten Sätze wurden in einem Tabellendokument (Open Office) dargestellt. Hierzu wurde der Laptop hp Compaq nc 6320 verwendet. Im Vorfeld wurde eine fortlaufende Tabelle angefertigt, die für jeden Probanden die richtigen Sätze in der richtigen Reihenfolge beinhaltete.

Die erste Spalte beinhaltete die Satznummer im jeweiligen Block und die darauffolgenden 8 Spalten waren für die einzelnen Wörter des Satzes vorgesehen.

Jedem Wort wurde von links beginnend genau ein Feld zugeteilt, so dass bei allen Sätzen, deren Wortanzahl weniger als 8 betrug, Felder frei blieben. Hinter jedem Satz gab es weitere 8 kleine Felder, die der weiteren Auswertung dienten.

Um eine schnelle mitprotokollierende Erfassung der Probandenantworten zu ermöglichen, wurde ein Makro programmiert. Für eine richtig gegebene Antwort eines Probanden wurde eine Taste definiert, für jede falsche Antwort eine andere Taste. Indem der Versuchsleiter die Taste für die richtige Antwort betätigte, wurde das Tabellenfeld, in dem das Wort stand, grün eingefärbt und ebenfalls das korrespondierende kleine Feld, das sich den Wortspalten anschloss. Außerdem sprang der Cursor automatisch ein Feld weiter. Bei einer falsch gegebenen Antwort wurden analog dazu die beiden korrespondierenden Felder orange eingefärbt. So war es möglich jedem Wort die Angabe richtig oder falsch zuzuweisen. Zur Korrektur für ein falsch markiertes Feld gab es eine weitere Taste, die das jeweilige Feld wiederum entfärbte. Eine weitere Taste ermöglichte den Zeilenwechsel. Außerdem war es möglich sich mit den Navigationstasten in der

Tabelle 2.2: Blockverteilung auf die Probanden

	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block10	Block11	Block12
Proband												
1	x	x	x	x								
2		x	x	x	x							
3			x	x	x	x						
4				x	x	x	x					
5					x	x	x	x				
6						x	x	x	x			
7							x	x	x	x		
8								x	x	x	x	
9									x	x	x	x
10	x									x	x	x
11	x	x									x	x
12	x	x	x									x
13	x	x	x	x								
14		x	x	x	x							
15			x	x	x	x						
16				x	x	x	x					
17					x	x	x	x				
18						x	x	x	x			
19							x	x	x	x		
20								x	x	x	x	
21									x	x	x	x
22	x									x	x	x
23	x	x									x	x
24	x	x	x									x
25	x	x	x	x								
26		x	x	x	x							
27			x	x	x	x						
28				x	x	x	x					
29					x	x	x	x				
30						x	x	x	x			
31							x	x	x	x		
32								x	x	x	x	
33									x	x	x	x
34	x									x	x	x
35	x	x									x	x
36	x	x	x									x
37		x	x	x	x							

Tabelle zu bewegen. Zur Veranschaulichung der beschriebenen Vorgehensweise ist ein Ausschnitt dieser Tabelle in Abbildung 2.3 auf S. 17 dargestellt.

So war nach Abschluss aller Messungen eine Tabellendatei vorhanden, die alle Versuchsdaten beinhaltet. Zur weiteren Auswertung wurden diese Daten so exportiert, dass sie vom Statistik-Programm „R“ eingelesen werden konnten. Dieses Programm bildete die Grundlage bei der Auswertung der Ergebnisse.

2.4 Methode der Auswertung für die Listenzusammenstellung

Für das Zusammenstellen der Listen wurde mit dem Statistikprogramm „R“ ein Auswertungsskript verfasst. Dabei gab es vorab Bedingungen, die durch die Struktur der Listen festgelegt waren. Demnach beinhaltete jede Liste 10 Sätze mit folgender Zuordnung:

- 3-Wort-Satz: Aussage
- 4-Wort-Satz: Ausruf oder Kommando
- 4-Wort-Satz: Frage
- 5-Wort-Satz: Ausruf oder Kommando
- 5-Wort-Satz: Frage
- 5-Wort-Satz: Aussage
- 6-Wort-Satz: Frage
- 6-Wort-Satz: Aussage
- 7-Wort-Satz: Aussage
- 8-Wort-Satz: Aussage

Die Reihenfolge der endgültigen Satzlisten orientierte sich an den Vorgaben des originalen HSM-Satztests. So resultierten in jeder Liste 10 verschiedene Satzarten (Satzkategorie und Wortanzahl). Die Sätze jeder Satzart wurden jeweils per Zufall auf alle Listen verteilt. Nun konnte der durchschnittliche Verständniswert jeder Liste berechnet werden, indem die korrekt verstandenen Worte der jeweiligen Liste in Bezug gesetzt wurden zur Gesamtanzahl der Worte jeder Liste. Dazu wurden als erstes alle verstandenen Worte aller 12 Probanden, die denselben Satz gehört hatten, addiert. Dies wurde für alle Sätze einer Liste durchgeführt. Anschließend wurde die Wortanzahl aller Sätze in einer Liste addiert und mit 12 multipliziert. So erhielt man die maximal mögliche Anzahl korrekter Antworten bzw. Wörter für eine Liste. Nun konnte man die Anzahl aller

korrekt erkannten Wörter durch die maximal mögliche Anzahl der Wörter teilen und erhielt den Verständniswert der jeweiligen Liste. Multiplizierte man diesen Wert mit 100 erhielt man den Verständniswert einer Liste in Prozent. So ließ sich auch der mittlere Verständniswert aller Listen berechnen. Für einige Satzarten waren mehr Sätze vorhanden als benötigt wurden. Dies führte dazu, dass ein Restpool für die jeweilige Satzart übrigblieb. Durch den bestehenden Restpool gab es die Möglichkeit, dass verschiedene Zufallsverteilungen unterschiedliche mittlere Verständniswerte lieferten.

Die Listen wurden nun einem Austauschprozess unterzogen. Ziel dieses Prozesses war es, Listen exakt gleicher Schwierigkeit zu erhalten. Für den Austauschprozess wurden aber nur solche initiale Zufalls-Listenzuteilungen berücksichtigt, deren Gesamtanzahl verstandener Worte dividiert durch die Anzahl aller Listen eine ganze Zahl ergab. Denn nur so war es überhaupt möglich, eine Listenzuteilung durchzuführen, die am Ende zu lauter Listen gleicher Schwierigkeit führte.

Nach der Zufalls-Listenzuordnung wurden die beiden Listen, deren Verständniswerte am meisten voneinander abwichen, ausgewählt und durch Austausch von Sätzen gleicher Satzart eine Angleichung des Verständniswertes versucht. Dabei waren theoretisch $2^{10} = 1024$ Kombinationen möglich. Durch die Gleichwertigkeit der Listen handelte es sich aber nur um 512 *verschiedene* Kombinationen. Alle 512 Verständniswerte wurden berechnet und diejenige Kombination akzeptiert, die den geringsten Verständniswertunterschied beider Listen aufwies. Bei einem Verständniswertunterschied von 0 wurde ggf. früher abgebrochen. Nach dem Austauschprozess wurden die nächsten beiden Listen ausgewählt, die nun die größte Differenz des Verständniswertes aufwiesen. Auf diese Art und Weise wurde repetitiv fortgefahren bis kein Satzaustausch zwischen zwei Listen mehr zu einer Verbesserung führte.

Als Ergebnis dieses Austauschprozesses erhielt man lauter Listen gleicher Schwierigkeit.

3 Ergebnisse

3.1 Auswertungsdaten

Insgesamt 16656 Datensätze – 1388 Sätze zu je 12 Durchläufen – bildeten die Grundlage der Auswertung. Diese wurde mit dem Statistikprogramm „R“ (Sachs und Hedderich, 2009) durchgeführt.

So war es möglich verschiedene Aspekte zu betrachten, auf die im Folgenden genauer eingegangen wird.

3.2 Häufigkeitsverteilung der Sätze nach Verständniswert

Basierend auf die dichotomen Daten der Einzelwörter, d. h. „verstanden“ oder „nicht verstanden“, konnte der Verständniswert jedes Satzes errechnet werden (s. Abschnitt 2.4 auf S. 15). Multipliziert man diesen mit 100, so ergibt sich der prozentuale Verständniswert. Die Häufigkeitsverteilung der Sätze nach prozentualem Verständniswert ist in einem Histogramm (s. Abbildung 3.1 auf S. 19) graphisch dargestellt.

Für Sätze, die zu 0 % oder zu 100 % verstanden wurden, wurde jeweils ein eigener Balken reserviert, während für die restlichen Prozentwerte größere Klassenbreiten zu einem Balken zusammengefasst wurden. Es zeigt sich, dass nicht oder kaum verstandene Sätze etwas häufiger vorkommen, wohingegen bei gutem bis komplettem Verständnis die Häufigkeit abnimmt.

Weiterhin wurde die Verteilung der Sätze nach Verständniswert in Abhängigkeit von der jeweiligen Satzlänge (3- bis 8-Wort-Sätze) untersucht. Die Ergebnisse sind in den Histogrammen 3.2, 3.3, 3.4, 3.5, 3.6, 3.7 graphisch dargestellt.

Tendenziell lässt sich das Verteilungsmuster, das für alle Sätze zusammen gilt, auch bei der Betrachtung der einzelnen Satzlengthen wiederfinden.

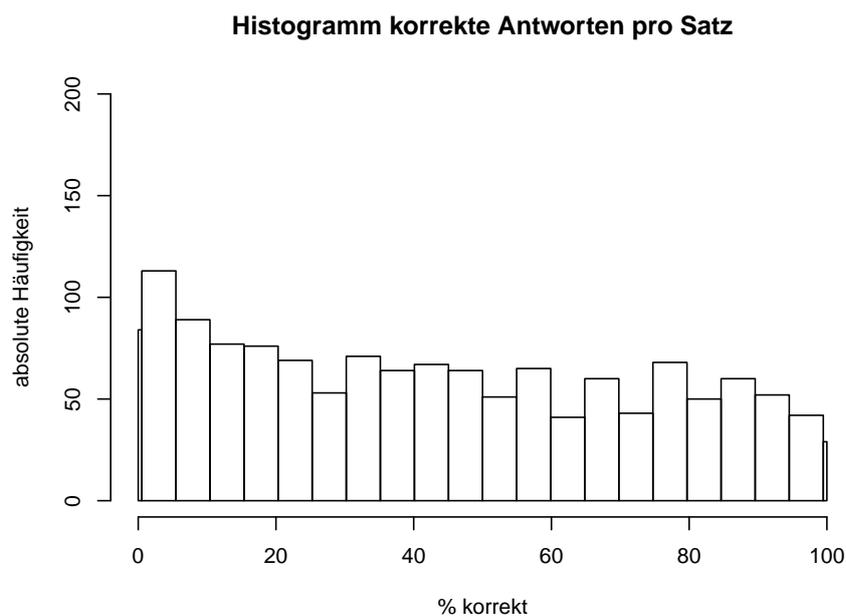


Abbildung 3.1: Häufigkeitsverteilung aller Sätze nach Verständniswert

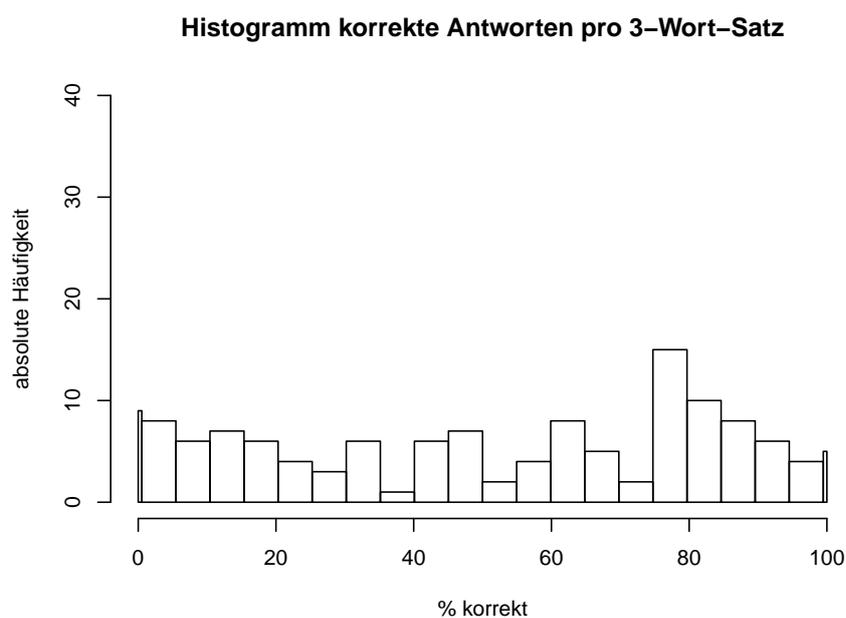


Abbildung 3.2: Häufigkeitsverteilung der 3-Wort-Sätze nach Verständniswert

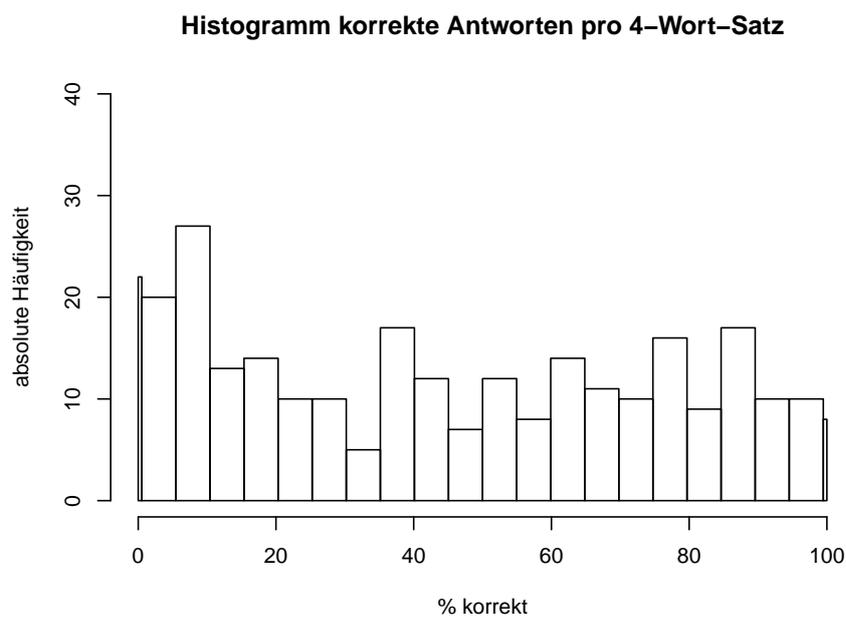


Abbildung 3.3: Häufigkeitsverteilung der 4-Wortsätze nach Verständniswert

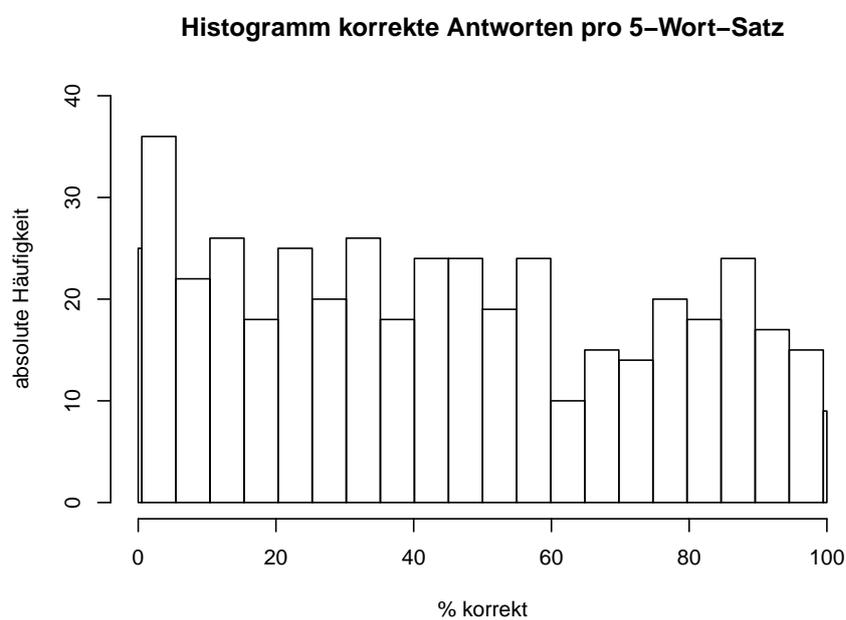


Abbildung 3.4: Häufigkeitsverteilung der 5-Wortsätze nach Verständniswert

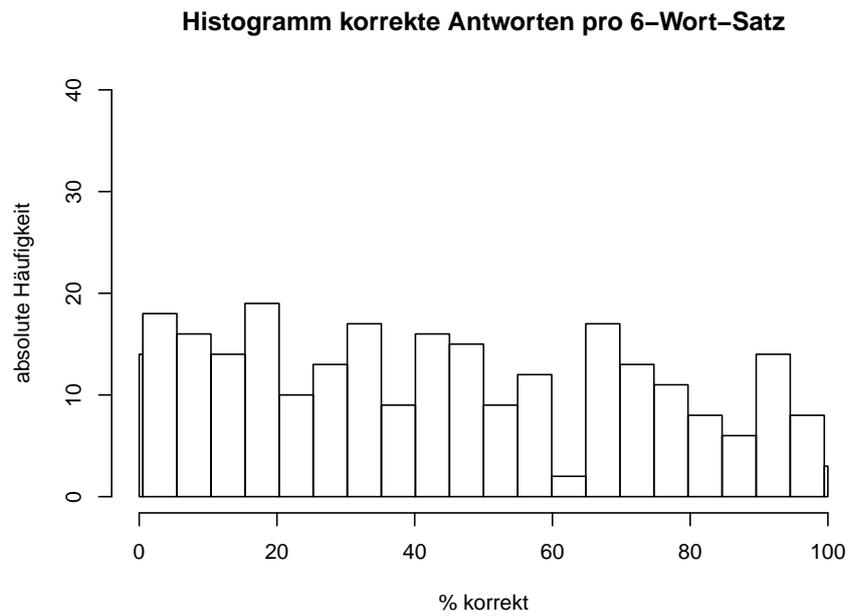


Abbildung 3.5: Häufigkeitsverteilung der 6-Wortsätze nach Verständniswert

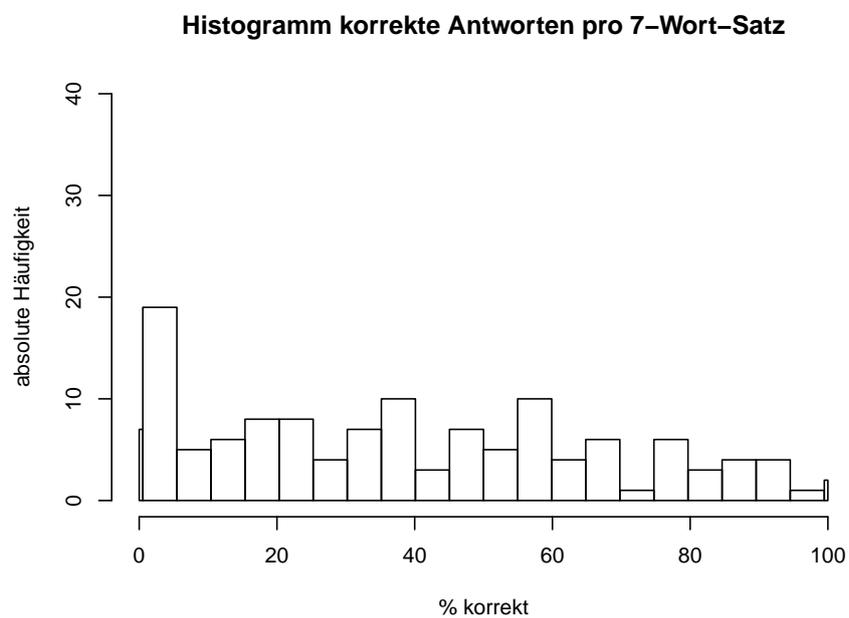


Abbildung 3.6: Häufigkeitsverteilung der 7-Wortsätze nach Verständniswert

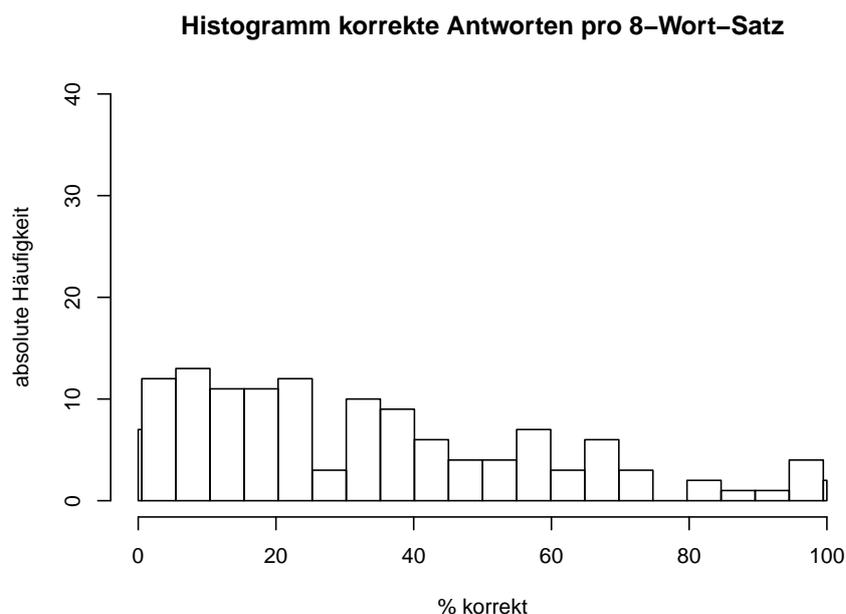


Abbildung 3.7: Häufigkeitsverteilung der 8-Wortsätze nach Verständniswert

3.3 Einflussfaktoren auf das Satzverständnis

Im Folgenden wurde der Verständniswert der Sätze in Abhängigkeit von verschiedenen Faktoren untersucht. Die Daten wurden auf Ordinalskalenniveau behandelt. Eine Behandlung auf metrischem Skalenniveau war nicht angebracht, da Häufigkeitswerte zu Grunde lagen. Zur statistischen Evaluierung konnte der Kruskal-Wallis-Test als globaler multipler Test und der Wilcoxon-Test als Einzeltest verwendet werden.

Für die graphische Darstellung wurden Boxplots gewählt. Ein Boxplot enthält alle, also „n“ Urteile, die in die jeweilige Kategorie fallen. Folgende Werte können dem Boxplot entnommen werden:

- Mittelwert: schwarzer Punkt
- Median: dicker schwarzer Querstrich
- 1. Quartil: untere Kästchenbegrenzung
- 3. Quartil: obere Kästchenbegrenzung
- Minimalwert: unterer dünner Strich
- Maximalwert: oberer dünner Strich

3.3.1 Abhängigkeit von der Satzlänge

Die Verständniswerte aller Sätze gleicher Länge bildeten die jeweilige Datenmenge. Der Zusammenhang zwischen Satzlänge und Satzverständniswert ist in Abbildung 3.8 auf S. 24 dargestellt. Der Kruskal-Wallis-Test ergab ein $p < 1,0 \times 10^{-6}$, es besteht also eine hochsignifikante Abhängigkeit des Verständniswertes von der Satzlänge.

Im originalen HSM-Satztest bestand eine Testeinheit aus zwei Listen zu je 10 Sätzen. Dabei war die Anzahl der Sätze für jede Satzlänge und weiterhin die genaue Satzkategorie für jeden Satz festgelegt. Auf dieser Koppelung der Satzlänge mit der Satzkategorie basiert auch die Vorarbeit von Angermüller (unveröff. Diss.). Um einen möglichen Einfluss der Satzkategorien auszuschließen, wurde in einer Zusatzbetrachtung nur die Teilmenge der Aussagesätze herausgegriffen. Anhand dieser Daten konnte der direkte Zusammenhang zwischen Satzlänge und Verständniswert graphisch in Abbildung 3.9 auf S. 24 dargestellt werden. Der Kruskal-Wallis-Tests zeigt mit $p < 1,0 \times 10^{-6}$ ein hohes Signifikanzniveau. Demzufolge existiert der Einfluss der Satzlängen unabhängig von der Satzkategorie.

3.3.2 Abhängigkeit von der Satzkategorie

Eine weitere Betrachtung bezog sich auf die Abhängigkeit des Satzverständnisses von der Satzkategorie. Wie schon erwähnt gab es folgende Satzkategorien: Aussage, Frage, Ausruf, Kommando. Die Verständniswerte aller Sätze wurden getrennt nach Satzkategorie betrachtet. Mit Hilfe dieser Daten wurde ein Boxplot (s. Abbildung 3.10 auf S. 26) erstellt. Das Ergebnis des Kruskal-Wallis-Tests zeigt eine hohe Signifikanz mit $p < 1,0 \times 10^{-6}$ bezüglich der Abhängigkeit des Verständnisses von der Satzkategorie. Weiterhin führten wir Paarvergleiche der einzelnen Satzkategorien mit dem Wilcoxon-Test durch (s. Tabelle 3.1 auf S. 25).

Es zeigt sich, dass für den Paarvergleich Frage - Ausruf kein signifikanter Zusammenhang besteht. Für alle anderen Paarvergleiche liegt ein signifikanter Zusammenhang auf dem 1-Prozent-Niveau vor. Wie oben erwähnt sind Satzlänge und Satzkategorie miteinander gekoppelt, so dass auch hier die Betrachtung präzisiert werden musste. Deshalb wurden nun Sätze mit unterschiedlicher Satzkategorie, aber gleicher Satzlänge betrachtet. So kann der Einfluss der Satzlän-

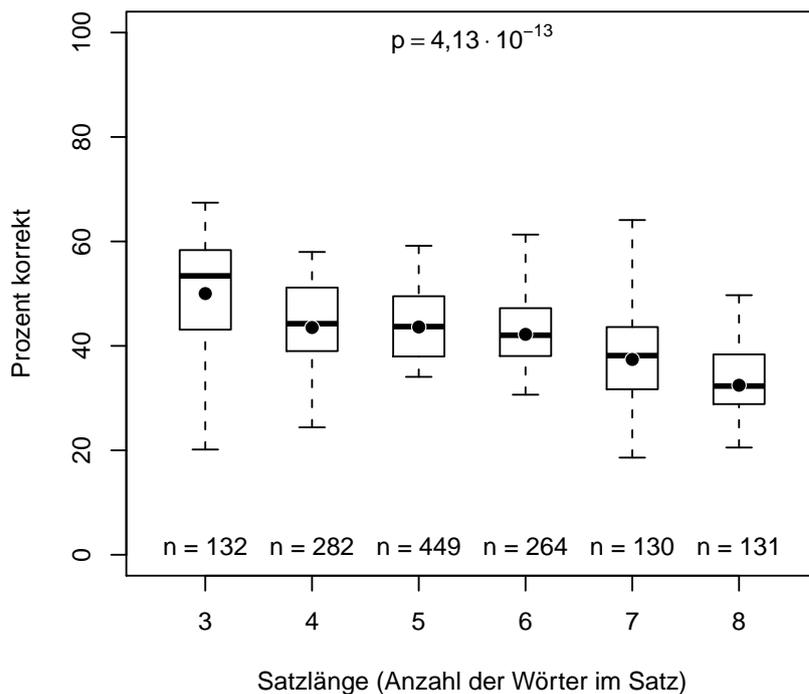


Abbildung 3.8: Abhängigkeit des Satzverständnisses von der Satzlänge
 n = Anzahl der Sätze pro Boxplot; p = Signifikanz

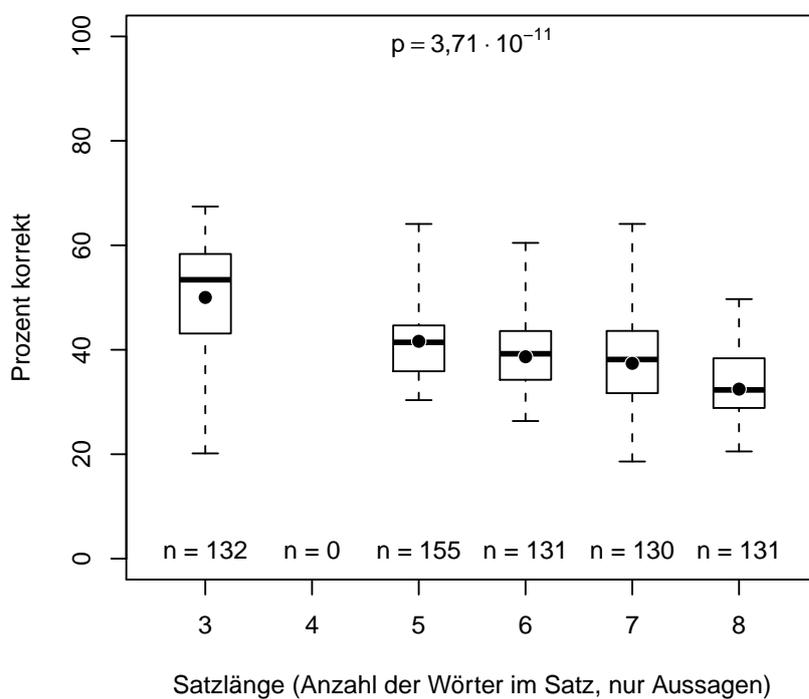


Abbildung 3.9: Abhängigkeit des Satzverständnisses von der Satzlänge
 n = Anzahl der Sätze pro Boxplot; p = Signifikanz

Tabelle 3.1: Paarvergleich (Wilcoxon-Test) der Satzkategorien

Satzkategorie			p-Wert
Aussage	—	Frage	$< 1,0 \times 10^{-6}$
Aussage	—	Ausruf	$< 1,0 \times 10^{-6}$
Aussage	—	Befehl	0,000991
Frage	—	Ausruf	0,126644
Frage	—	Befehl	$< 1,0 \times 10^{-6}$
Ausruf	—	Befehl	$< 1,0 \times 10^{-6}$

Tabelle 3.2: Paarvergleich (Wilcoxon-Test) der Satzkategorien für 5-Wortsätze

Satzkategorie			p-Wert
Aussage	—	Frage	0,001057
Aussage	—	Ausruf	0,000002
Aussage	—	Befehl	0,000500
Frage	—	Ausruf	0,107946
Frage	—	Befehl	$< 1,0 \times 10^{-6}$
Ausruf	—	Befehl	0,000002

ge ausgeschaltet werden. Diese Bedingung fand sich nur beim 5-Wort-Satz. Mit Hilfe dieser Daten wurde ein Boxplot (s. Abbildung 3.11 auf S. 26) erstellt, der den Zusammenhang graphisch zeigt.

Der Kruskal-Wallis-Test ergibt mit $p < 1,0 \times 10^{-6}$ einen hochsignifikanten Zusammenhang. Auch für diese Betrachtung – Satzkategorie ohne Einfluss der Satzlänge – führten wir Paarvergleiche durch (s. Tabelle 3.2 auf S. 25).

Hier zeigt sich dasselbe Muster wie bei der Berechnung der Signifikanzen mit Einfluss der Satzlänge. Lediglich die einzelnen Werte sind minimal niedriger. Dies war wohl auf die kleinere Anzahl an Items zurückzuführen. Sicherlich sind Aussagen und Fragen bei beiden Betrachtungen gut vergleichbar, da genügend Einzelwerte zur Berechnung vorlagen. Für Ausrufe und Befehle muss man anmerken, dass diese beiden Satzkategorien nur für den hier untersuchten 5-Wort-Satz und zusätzlich für den 4-Wort-Satz vorhanden waren. Da nun die Verständnismerte der 4- und 5-Wort-Sätze graphisch sehr ähnlich waren (s. Abbildung 3.8 auf S. 24), war bei der alleinigen Betrachtung der 5-Wort-Sätze kein Unterschied zu erwarten.

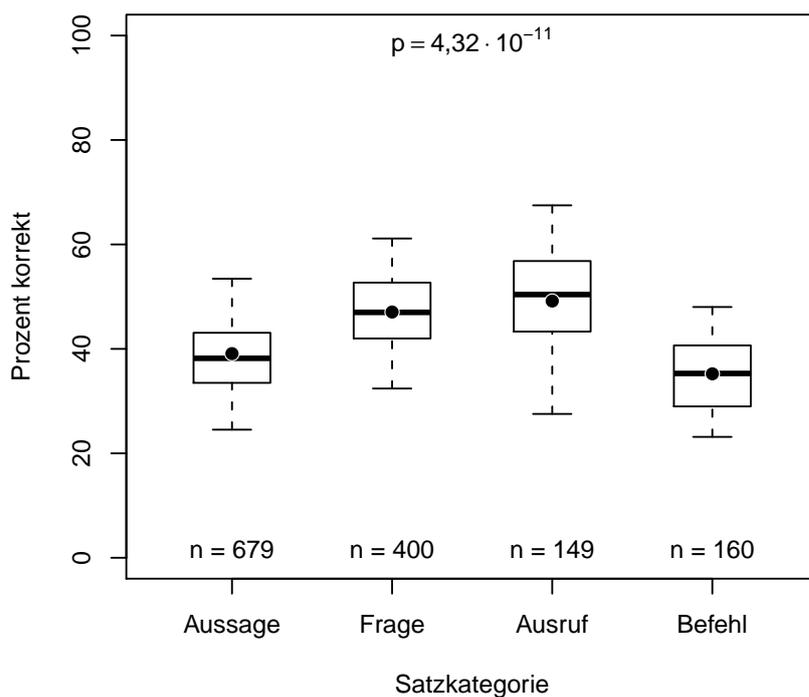


Abbildung 3.10: Abhängigkeit des Satzverständnisses von der Satzkatgorie
n = Anzahl der Sätze; p = Signifikanz

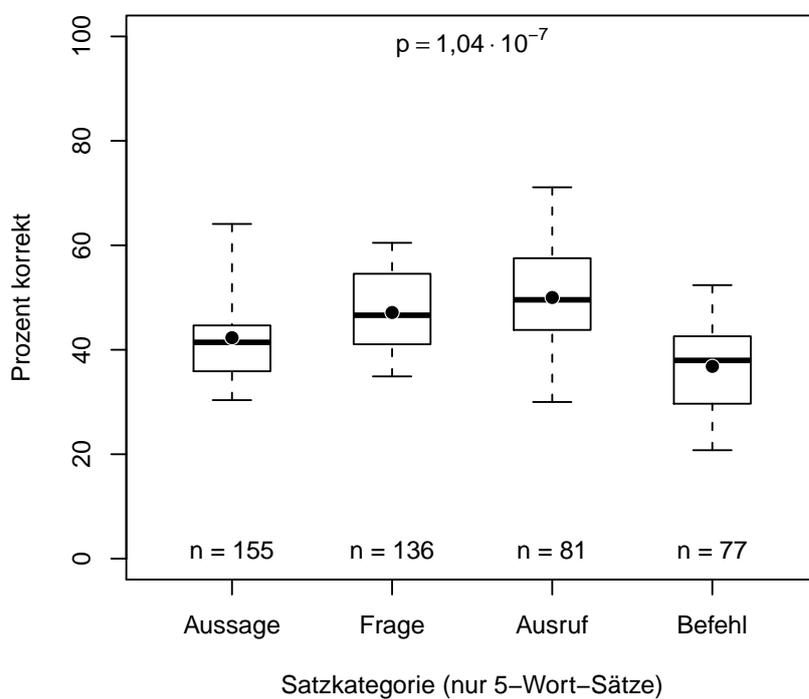


Abbildung 3.11: Abhängigkeit des Satzverständnisses von der Satzkatgorie
n = Anzahl der Sätze; p = Signifikanz

Tabelle 3.3: Paarvergleich (Wilcoxon-Test) der Zeitpositionen

Zeitposition	p-Wert
01 — 02	0,000030
01 — 03	0,000030
01 — 04	$< 1,0 \times 10^{-6}$
02 — 03	0,330293
02 — 04	0,082547
03 — 04	0,635737

3.3.3 Abhängigkeit von der Zeitposition

Weiterhin war die mögliche Abhängigkeit des Verständniwertes von der Zeitposition von Interesse. Als Zeitposition wurde Durchlauf 1 bis 4 unterschieden. Jeder Proband bekam 4 der 12 Satzblöcke als Testmaterial zugewiesen (s. Abschnitt 2.3.2 auf S. 12). Es interessierte, ob die Satzblöcke unterschiedlich gut verstanden wurden, je nachdem zu welchem Zeitpunkt sie dem Probanden präsentiert wurden. Dabei war der erste Block derjenige, der dem Probanden beim ersten Durchlauf dargeboten wurde und der vierte Block derjenige, der dem Probanden im vierten Durchlauf dargeboten wurde. Jeder Block wurde an jeder Zeitposition gleich oft, nämlich 3 mal, abgehört (s. Abbildung 2.2 auf S. 14). Da es 12 Satzblöcke gab, bildeten für die jeweiligen Zeitpositionen 36 Werte die Grundlage der Berechnung. Die Ergebnisse sind graphisch in einem Boxplot (s. Abbildung 3.12 auf S. 28) dargestellt. Der Kruskal-Wallis-Test ergibt ein p von 0,010678 und damit eine Signifikanz auf dem 5%-Niveau, nicht jedoch auf dem 1%-Niveau. Weiterhin wurden Paarvergleiche für jeweils zwei Zeitpositionen mit dem Wilcoxon-Test durchgeführt (s. Tabelle 3.3 auf S. 27).

Beim Vergleich von Zeitposition 1 mit allen anderen Zeitpositionen zeigt sich eine Signifikanz auf dem 1-Prozent-Niveau. Während sich die anderen Zeitpositionen in ihrem Signifikanzniveau untereinander nicht unterschieden.

3.3.4 Abhängigkeit von der Formelhaftigkeit

Inwiefern Sprichwörter und Redewendungen - also die Formelhaftigkeit eines Satzes - tatsächlich den Verständnisswert beeinflussen, haben wir an dieser Stelle untersucht. Dazu wurden drei Kategorien der Formelhaftigkeit unterschieden: gering – mäßig – stark. Die Einteilung aller Sätze in eine der drei Kategorien

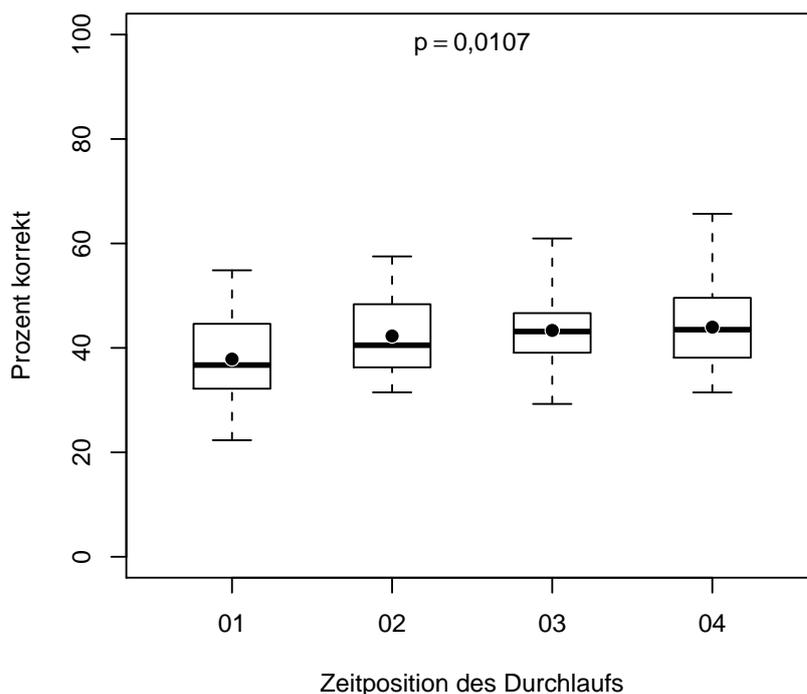


Abbildung 3.12: Abhängigkeit des Satzverständnisses von der Zeitposition
p = Signifikanz

wurde vom Verfasser der Dissertation durchgeführt. Er erfüllte alle Kriterien, die an die teilnehmenden Probanden gestellt wurden. Somit kann davon ausgegangen werden, dass auch die Probanden subjektiv gesehen, diese oder eine sehr ähnliche Einteilung der Sätze zu den drei unterschiedlichen Kategorien der Formelhaftigkeit vorgenommen hätten.

Der errechnete Zusammenhang zwischen Formelhaftigkeit und Verständniswert ist graphisch in einem weiteren Boxplot (s. Abbildung 3.13 auf S. 29) dargestellt. Der Kruskal-Wallis-Tests ergibt ein $p = 0,000018$ und damit eine hohe Signifikanz. Es ist zu erwähnen, dass sich die Anzahl der Sätze in den drei Kategorien sehr stark unterscheidet, so dass die Aussagekraft der Statistik kritisch gesehen werden kann.

3.3.5 Abhängigkeit vom Geschlecht der Probanden

Die nächste Betrachtung befasst sich mit der Abhängigkeit des Verständniswertes vom Geschlecht der Probanden. Die Anzahl der weiblichen und männlichen Probanden war gleich; am Versuch nahmen je 18 weibliche und 18 männliche Probanden teil. Graphisch wird der betrachtete Zusammenhang durch einen

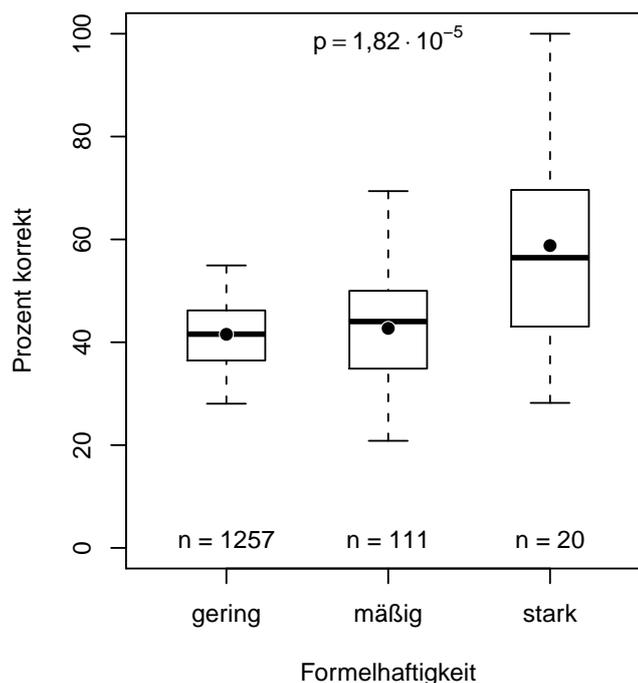


Abbildung 3.13: Abhängigkeit des Satzverständnisses von der Formelhaftigkeit
 n = Anzahl der Sätze; p = Signifikanz

Boxplot in Abbildung 3.14 auf S. 30 dargestellt. Der Wilcoxon-Test ergibt eine Signifikanz von $p = 0,401793$. Es liegt also keine signifikante Abhängigkeit des Verständniswertes vom Geschlecht vor.

3.3.6 Abhängigkeit vom Inhalt der Satzblöcke

Eine weitere Betrachtung befasste sich mit der Abhängigkeit des Verständniswertes vom Inhalt der Satzblöcke. Es wurden die Unterschiede zwischen den einzelnen Blöcken bzgl. des Verständniswertes untersucht. Es gab 12 verschiedene Blöcke zu je 115 oder 116 Wörtern. Jedem Probanden wurden vier der 12 Blöcke zugeteilt (s. Abschnitt 2.3.2 auf S. 12). So wurde jeder Block 12 mal abgehört (s. Abbildung 2.2 auf S. 14). Die Sätze desselben Blockes wurden jedes Mal in einer anderen, zufälligen Reihenfolge dargeboten. Der Zusammenhang zwischen Verständniswert und Blockinhalt bzw. Blocknummer ist mit Hilfe einer Boxplotgraphik in Abbildung 3.15 auf S. 30 dargestellt. Der Kruskal-Wallis-Test ergibt eine Signifikanz von $p = 0,000003$ und damit eine hochsignifikante Abhängigkeit des Verständnisses vom dargebotenen Block.

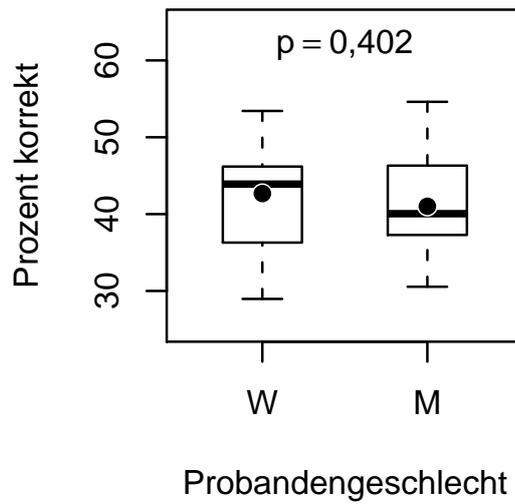


Abbildung 3.14: Abhängigkeit des Satzverständnisses vom Geschlecht
 $p = \text{Signifikanz}$

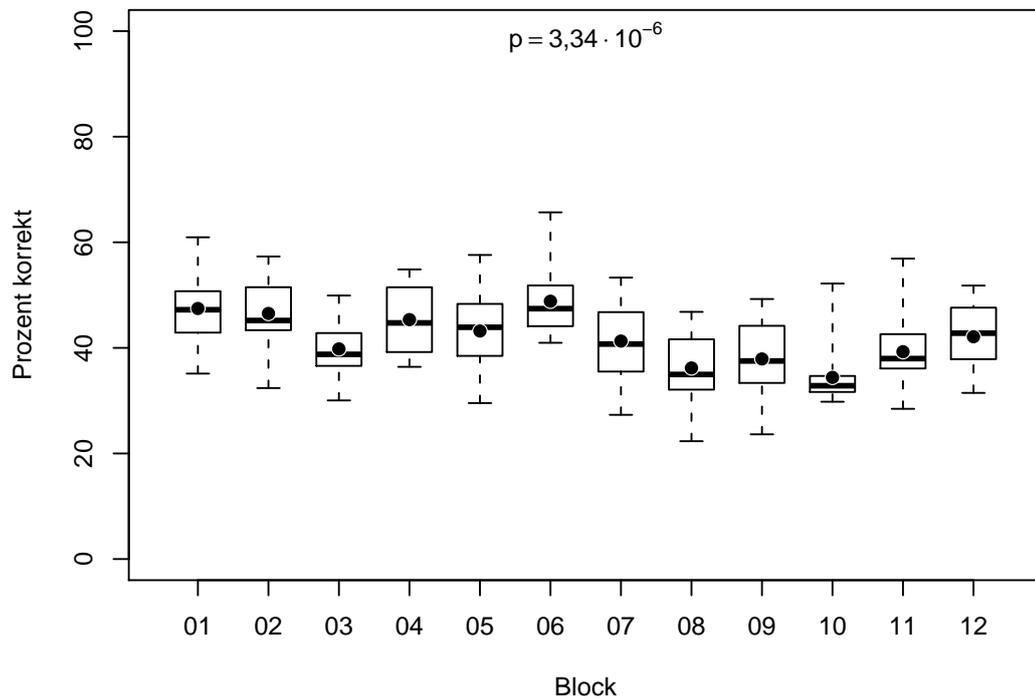


Abbildung 3.15: Abhängigkeit des Satzverständnisses vom Blockinhalt
 $p = \text{Signifikanz}$

3.4 Listenzusammenstellung

Ziel der Arbeit war es aus einem erweiterten Satzpool mehr Listen als beim originalen HSM-Satztest zusammenzustellen, die wiederum untereinander gleich schwierig sein sollten. Als erstes musste dazu die Schwierigkeit, das heißt der Verständnismesswert erhoben und anhand dessen und weiteren Vorgaben die Erstellung der Listen durchgeführt werden. Die genaue Methode hierfür wurde im Methodenkapitel erläutert (s. Abschnitt 2.4 auf S. 15). Als Ergebnis erhielten wir 130 exakt gleich schwere Listen für eine initiale Zufalls-Listenzuteilung.

Das Hauptkriterium der Vergleichbarkeit von Listen war ein übereinstimmender Verständnismesswert. Die Gleichheit aller Listen bzgl. dieses Kriteriums wurde mit der beschriebenen Methode erreicht. Ein weitergehendes Kriterium war die Varianz der Schwierigkeitsgrade der einzelnen Sätze innerhalb einer Liste. So könnten beispielsweise in einer Liste alle Sätze ungefähr zur Hälfte verstanden werden oder die Hälfte der Sätze komplett und die andere Hälfte gar nicht verstanden werden. Deshalb wurden 10000 Zufalls-Listenzuteilungen, die jeweils gleich schwere Listen lieferten, näher betrachtet. Es wurden jeweils sowohl die mittlere absolute Abweichung (MAD) als auch die Standardabweichung (SD) der Sätze jeder Liste einer Zufalls-Listenzuteilung berechnet. So erhielten wir für jede der 10000 Zufalls-Listenzuteilungen jeweils 130 MAD-Werte und 130 SD-Werte. Die Differenz zwischen dem größten und dem kleinsten der jeweils 130 MAD- bzw. SD- Werte (also die Spannweite) war ein Maß dafür, wie weit sich die beiden extremsten Listen in ihrer Varianz unterschieden. Der Mittelwert aller MAD bzw. SD-Werte der 130 Listen galt als Maß für die durchschnittliche Varianz einer Zufalls-Listenzuteilung. Um festzustellen, ob eine Korrelation zwischen dem Mittelwert und der Spannweite besteht, wurden die jeweiligen Werte für die 10000 Zufalls-Listenzuteilung graphisch dargestellt (s. Abbildung 3.16 auf S. 32). Da die MAD-Werte als das robustere Maß angesehen werden, wurde die Graphik anhand dieser Werte erstellt. Es war keine Korrelation feststellbar. Die Werte für die Spannweite erstreckten sich über eine Spanne von 0,3, während die Werte für die Mittelwerte eine Spanne von 0,05 einnahmen. Aufgrund dieser Werte und mit dem Ziel möglichst homogene Listen zu erreichen, wurde die Spannweite als das schwerwiegendere Kriterium erachtet. Die Zufalls-Listenzuteilung mit der unter den genannten Voraussetzungen niedrigsten Varianz hatte eine Spannweite der MAD-Werte von 0,380063 und einen Mittelwert der MAD-Werte von 0,351450. Die Listenzuteilung mit diesen Werten ist in der Graphik mit einem Kreis mar-

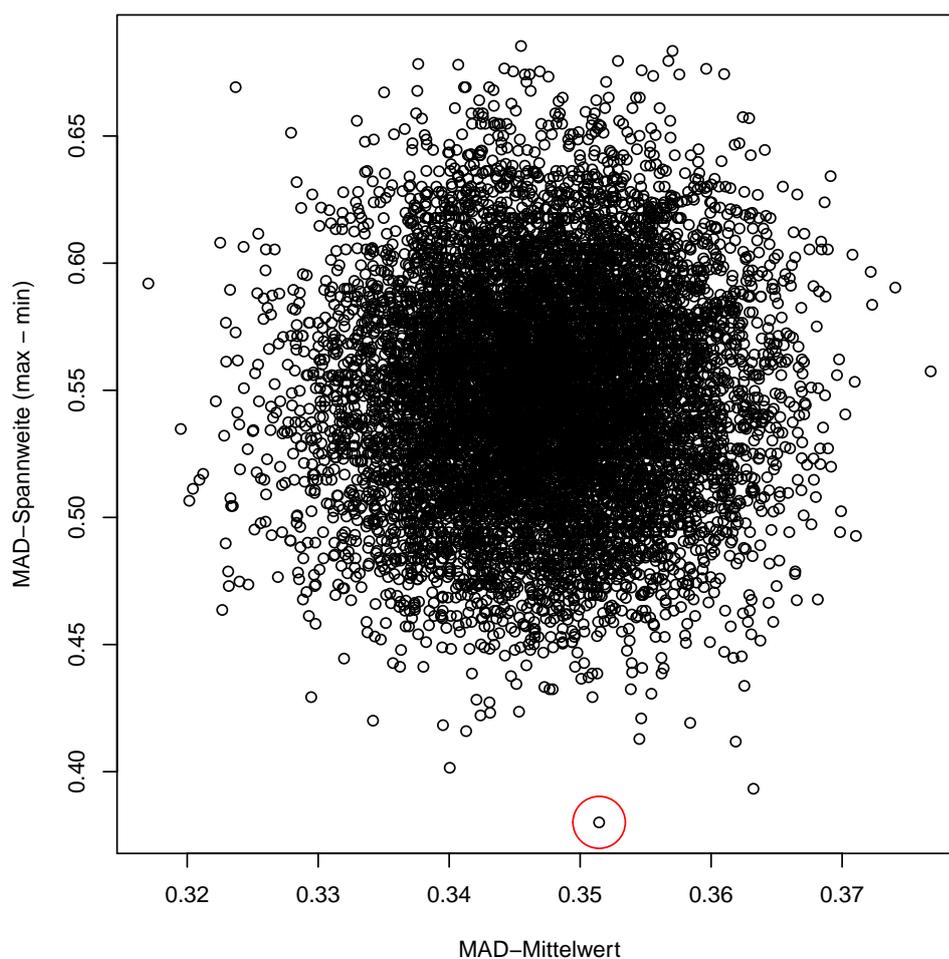


Abbildung 3.16: Korrelation zwischen Streubreite und Mittelwert
MAD-Diff = Spannweite der MAD-Werte
MAD-Mittelwert = Mittelwerte der MAD-Werte

kiert (s. Abbildung 3.16 auf S. 32). So gelang es schlussendlich dem Kriterium – gleicher Schwierigkeitsgrad für alle Listen – gerecht zu werden und darüber hinaus noch das Kriterium der Varianz innerhalb der Listen zu beachten.

4 Diskussion

4.1 Überblick

Die neu aufgesprochenen originalen HSM-Sätze und die hinzugefügten Sätze wurden mit Rauschen versehen und von den Probanden abgehört. So konnte jedem Satz ein durchschnittlicher Schwierigkeitswert zugeordnet werden. Anschließend wurden diese Sätze in Listen eingeteilt, so dass schlussendlich jede Liste genau gleich schwierig und damit austauschbar war. Somit liegt ein besseres Ergebnis vor, als bei dem originalen HSM-Satztest. Hier differierte der Verständniswert der einzelnen Satzgruppen zwischen 49,8 Prozent und 52,6 Prozent (Hochmair-Desoyer et al., 1997). Zum anderen konnte durch die Aufstockung des Testmaterials dem mehrfach geäußerten Kritikpunkt der Redundanz durch Wiederholungseffekte (Mrowinski und Scholz, 2006; Müller-Deile, 2009) entgegengetreten werden.

4.2 Probandenklientel

Es ist festzuhalten, dass die Ergebnisse und damit die Zusammenstellung der Satzlisten auf einer homogenen Probandenklientel basieren. Diese Probandengruppe bestand aus Studenten ausschließlich deutscher Muttersprache. Somit ist eine gute Belastbarkeit, eine gewisse intellektuelle Fähigkeit und damit auch eine hohe Sprachkompetenz anzunehmen: psychologische wie auch linguistische und phonetische Faktoren haben Einfluss auf das Sprachverstehen (Lindgren und Lindblom, 1983). Die Ergebnisse dieser Probandengruppe dürften somit über denen des Bevölkerungsdurchschnitts liegen und daher nicht ohne weiteres auf diese übertragbar sein. Allerdings gelang es mit der vorliegenden Probandengruppe vermutlich eine geringere Streuung der Ergebnisse zu erreichen. Dies würde die durchzuführende Evaluation begünstigen.

4.3 Satzgruppen in anderen Arbeiten

Von einigen Autoren (Bangert, 1980; von Wedel, 1986) wird die Inhomogenität von Sprachmaterial kritisiert und die Ausgewogenheit und Vergleichbarkeit zwischen einzelnen Gruppen als Grundvoraussetzung für objektive Messungen angesehen.

Es gibt verschiedene Arbeiten, die sich mit der Vergleichbarkeit der Satzgruppen des originalen HSM-Tests bzw. des Innsbrucker Satztests befassen. Strohmaier (1998) untersuchte diese Fragestellung in seiner Dissertation anhand eines vergleichbaren Probandenkollektivs wie es in der hiesigen Arbeit vorliegt. Er führte eine Klasseneinteilung anhand der durchschnittlichen Fehlerzahl pro Gruppe durch und errechnete mit dem Chi-Quadrat-Anpassungs-Test, dass die Hypothese der Normalverteilung nicht verworfen werden konnte. Dies war Begründung genug, eine Vergleichbarkeit der Satzgruppen anzunehmen und daraus die Austauschbarkeit abzuleiten. Auch Grebe (2000) stellte bei Normalhörenden um die 50 Jahre eine Normalverteilung fest und rechtfertigte damit die Austauschbarkeit der Satzgruppen. In der Dissertation von Zimmermann (2000) wurden Hörgeräteträger mit dem HSM-Satztest getestet. Die Verständnisquoten hatten eine relativ starke Streuung; jedoch zeigten mehrere Signifikanzberechnungen keine signifikanten Unterschiede zwischen den Satzgruppen. Es blieb offen, ob derartige Unterschiede in den Verständnisquoten auf die Satzgruppen selber zurückzuführen waren oder diese Unterschiede in der sehr inhomogenen Klientel der Hörgeräteträger lag. Deshalb forderte Zimmermann (2000), die Ausgewogenheit der Satzgruppen unter stärker standardisierten Bedingungen genauer zu untersuchen, mit dem konkreten Vorschlag ein homogenes Kollektiv Normalhörender dafür heranzuziehen. Die vorliegende Arbeit ist diesen Forderungen gerecht geworden. Es konnte bei einem homogenen Probandenklientel eine absolute Gleichheit der Satzgruppen bzgl. des Schwierigkeitsgrades erzielt werden.

4.4 Satzverständlichkeit

4.4.1 Allgemein

Das Durchschnittsverständnis aller abgehörten Sätze lag bei ca. 42 %. Ein 50prozentiges Verstehen wäre wünschenswert gewesen, da an dieser Stelle die größ-

te Empfindlichkeit besteht und dieser Punkt als die Sprachverständlichkeitsschwelle angesehen wird. Das heißt, dass hier schon kleine Abweichungen bei der SNR relativ starke Verständlichkeitswertänderungen herbeiführen. Die resultierenden 42 % sind trotzdem ein akzeptables Ergebnis. Zur Ermittlung der zu verwendenden SNR wurden von Angermüller (unveröff. Diss.) Vorversuche mit vier Testpersonen durchgeführt. Diese wurden bei unterschiedlichen SNRs getestet und aufgrund der Ergebnisse wurde eine SNR von -10 dB als Standard für die durchzuführende Versuchsreihe festgelegt. Möglicherweise haben diese Testpersonen aufgrund ihrer Vertrautheit mit dieser Thematik überdurchschnittlich gut abgeschnitten, weshalb in der Folge das Satzverständnis der an dieser Arbeit teilnehmenden Probanden im Schnitt unter 50 % lag. Diese Tatsache schränkt aber die Vergleichbarkeit der Satzlisten untereinander nicht ein.

Bei der vorliegenden Versuchsdurchführung fiel auf, dass sehr viele Sätze überhaupt nicht oder komplett richtig verstanden wurden. Dies ist graphisch in Abbildung 3.1 auf S. 19 dargestellt. Diesen Effekt hat Strohmaier (1998) bei der Betrachtung der originalen HSM-Sätze ebenfalls beobachtet – in einer noch extremeren Form, als es unsere Ergebnisse zeigten. Strohmaier (1998) führte diese Tatsache auf die zentrale Informationsverarbeitung zurück, die insbesondere bei Satztests zum Zuge kommt, im Gegensatz zu Testverfahren mit einzelnen Wörtern oder Silben, die in keinem weiteren Sinnzusammenhang stehen. Mit der zentralen Informationsverarbeitung und deren Einfluss auf das Sprachverstehen befassten sich auch Schultz-Coulon (1973), Schlöndorff und Tegtmeier (1971) und Lindgren und Lindblom (1983). Dabei stellte Schultz-Coulon (1973) fest, dass das Sprachverständnis von psychologischen Faktoren wie Konzentration, Intelligenz und Affektlage abhängig ist. Schlöndorff und Tegtmeier (1971) argumentierten, dass zum Hörvermögen der zentrale Aufwand gehört und dieser sich individuell unterscheidet, so dass es manch einem leichter und manch einem schwerer fällt, einen gesprochenen Text zu verstehen. Lindgren und Lindblom (1983) sahen neben psychologischen Aspekten v. a. auch linguistische Gründe für die Beeinflussung der zentralen Sprachverarbeitung.

Es scheint keine einfache objektivierbare Erklärung zu geben, so dass es fraglich ist, ob man überhaupt verhindern kann, dass eine so große Anzahl an extremen Sätzen vorkommt. Bei den erhobenen Daten gab es mehr nicht verstandene Sätze als komplett verstandene Sätze. Somit könnte eine Veränderung des SNRs, der zu einem mittleren Satzverständnis von 50 % führen würde – also einer Anhebung des SNRs - sich die Anzahl der nicht verstandenen Sätze re-

duzieren und die Anzahl der komplett verstandenen Sätze zunehmen und somit eine gleichmäßigere Verteilung der extremen Sätze resultieren; die Anzahl dieser Extremsätze bliebe dabei vermutlich ähnlich.

4.4.2 Einfluss der Satzlänge

Betrachtet man das Verständnis des Satzes in Abhängigkeit von seiner Satzlänge (s. Abschnitt 3.3.1 auf S. 23), dann ist zu erkennen, dass der Median so wie der Mittelwert der Verständniswerte mit Anzahl der Wörter abnimmt. Betrachtet man das Signifikanzniveau, erhält man hochsignifikante Unterschiede. Niemeyer (1967) beschreibt zwei Effekte, die beim Satzverständnis eine Rolle spielen. Auf der einen Seite erhöht sich in Sätzen im Gegensatz zu einzelnen Wörtern die informelle Reserve, andererseits muss das Gehörte zu größeren Sinnzusammenhängen integriert werden. Er vergleicht hierbei in erster Linie Sätze im Allgemeinen mit einzelnen Wörtern. Diesbezüglich stellt er fest, dass normalerweise die informellen Reserven im Satz die Erschwernis durch größere Sinnzusammenhänge überwiegen. All diese Aussagen sind für Sätze ohne Rauschen, allerdings im Zusammenhang mit Schwerhörigkeit gemacht worden. Weiterhin ist Morton (1964) der Meinung, dass jeder Kontext, der die Wahrscheinlichkeit einer Worterkennung erhöht, die Verständnisschwelle herabsetzt.

Betrachtet man die in dem Versuch mit Rauschen abgehörten Sätze nach ihrer Wortanzahl, scheint dies nicht zuzutreffen. Niemeyer (1967) und Morton (1964) lassen durch ihre Aussagen vermuten, dass eigentlich mit Erhöhung der Wortanzahl das Verständnis des Satzes zunehmen müsste. Dies entspricht nicht der Beobachtung in der durchgeführten Versuchsreihe. Es ist wahrscheinlich, dass durch das Rauschen andere Voraussetzungen geschaffen wurden.

Die in dieser Arbeit gemachte Beobachtung deckt sich eher mit derjenigen von Zimmermann (2000), der die Fehleranzahl pro Satz in Abhängigkeit der Listenposition untersucht hat. Die Mittelwerte der Fehlerzahlen steigen dabei mit zunehmender Wortanzahl an. Aufgrund dieser Beobachtung wäre es in Zukunft interessant zu untersuchen, wie sich die Satzlänge auf das Satzverständnis sowohl bei Bedingungen ohne Rauschen als auch mit Rauschen bei verschiedenen starken Rauschpegeln auswirkt. Hutchinson (1989) weist darauf hin, dass neben akustischen auch linguistische Aspekte wie Länge und Komplexität eines Satzes eine Rolle beim Sprachverständnis spielen.

4.4.3 Einfluss der Satzkategorie

Die Satzkategorie ist ein weiterer Einflussfaktor bezüglich der Satzverständlichkeit (s. Abschnitt 3.3.2 auf S. 23). In Abbildung 3.10 auf S. 26 ist dieser Zusammenhang dargestellt. Es ist zu erkennen, dass Ausrufe und Fragen besser - und ähnlich gut - verstehbar sind als Kommandos und Aussagen. Dies wurde durch die Betrachtung der Signifikanzniveaus durch Paarvergleiche (s. Abbildung 3.2 auf S. 25) deutlich, indem man feststellen konnte, dass für den Vergleich Frage - Ausruf kein signifikanter Unterschied bestand, für alle anderen Vergleiche jedoch hoch signifikante Unterschiede.

Niemeyer und Beckmann (1962) stellten fest, dass beim Satzverständnis Sprachdynamik, Sprachmelodie, Sprachklangfärbung, Tempoänderungen und Sprachrhythmus eine entscheidende Rolle spielen.

Die Grundfrequenz bei Fragen steigt zum Ende hin an, während sie bei Ausrufen abnimmt (Domes, 2009). Auch kommt es zum Anstieg der Grundfrequenz bei Betonungen, die für Ausrufe und Fragen charakteristisch sind. Die genannten Aspekte würden erklären, warum gerade Fragen und Ausrufe besser verstehbar waren.

Dieser Tatsache wird in der vorliegenden Arbeit insofern Rechnung getragen, als dass alle Listen genau gleich viele Sätze der verschiedenen Kategorien beinhalten - gekoppelt an die Wortlänge. Beispielsweise würde ein 7-Wort-Satz, der per se eher schlechter verstanden wird, theoretisch durch die Kategorie „Fragesatz“ in der Verständlichkeit aufgewertet. Da nun aber jede Liste an jeder Position genau dieselbe Satzlänge und Satzkategorie aufweist, werden alle Listen gleichermaßen durch die Satzkategorie beeinflusst. Somit ist ein verfälschender Effekt auf die Schwierigkeitsgrade der Listen durch unterschiedliche Satzkategorien auszuschließen. Kommandos und Ausrufe werden in der Graphik einzeln aufgeführt. Diese beiden Kategorien werden aber bei der Listeneinteilung als eine Einheit verwendet. Dies ist insofern problematisch, als sie sich unterschiedlich auf das Satzverständnis auswirken. Diese Vorgehensweise wurde schon in der Dissertation von Angermüller (unveröff. Diss.), die die Grundlage dieser Arbeit bildet, so festgelegt.

4.4.4 Einfluss der Zeitposition

Es stellte sich die Frage, inwiefern der Zeitpunkt, an dem ein Satzblock im Versuchsverlauf abgehört wird, eine Rolle spielt. Bei der Untersuchung traten Unterschiede je nach Zeitposition auf. In Abbildung 3.12 auf S. 28 zeigt sich, dass sich mit zunehmender Zeitposition die Verständlichkeit der Sätze erhöht. Bei der Betrachtung der Paarvergleiche (s. Abbildung 3.3 auf S. 27) sieht man eindeutig, dass sich die Verständlichkeit von Block 1 hoch signifikant von den Verständlichkeiten aller anderen Blöcke unterscheidet, während sich für den Vergleich der anderen Blöcke (2-4) untereinander keine signifikanten Unterschiede ergeben. Dies weist darauf hin, dass die Probanden Zeit brauchten, um sich an die verrauschten Sätze zu gewöhnen. Würde man andererseits die Probanden durch eine zu ausgedehnte Versuchsdurchführung überfordern, würden die Ergebnisse wahrscheinlich wieder schlechter werden, da mit der Zeit die Konzentration nachlässt; denn diese beeinflusst die zentrale Informationsverarbeitung (Schultz-Coulon, 1973) und damit indirekt die Ergebnisse. Bei diesem Versuch war die Satzmenge pro Proband trotz der beträchtlichen Anzahl akzeptabel. Würde man aber mit einer anderen, weniger belastbaren Probandenklientel diesen erweiterten Test durchführen, könnte es möglicherweise schon frühzeitig zu einer Verschlechterung der Ergebnisse kommen, so dass der Umfang des Testmaterials reduziert werden müsste. Für den ursprünglichen HSM-Test sind nur 20 Sätze pro Testdurchgang vorgesehen, so dass das Problem der zu großen Datenmenge nicht auftreten wird.

Dem signifikanten Unterschied der Zeitposition ist insofern Rechnung getragen, als dass alle 12 Blöcke gleich oft an allen vier Positionen abgehört werden (s. Abbildung 2.2 auf S. 14). Jeder Proband hat anfangs einen Probekblock von 20 Sätzen abgehört, um sich an die Situation zu gewöhnen. Möglicherweise sollte man in Zukunft einen etwas umfangreicheren Probekblock darbieten, so dass der Proband beim Start des eigentlichen Versuchs von Anfang an höhere Verständnismwerte aufweist. Damit würde der Einfluss der Zeitposition verringert werden können.

4.4.5 Einfluss der Formelhaftigkeit

In dem ursprünglichen HSM-Satztest wurde absichtlich auf Sprichwörter und Redewendungen verzichtet. Allerdings wurden Sätze kreiert, die dadurch entstan-

den, dass Sprichwörter oder Redewendungen minimal verändert wurden, beispielsweise durch Austausch eines Wortes. Dies könnte dazu führen, dass es schwieriger ist diese Sätze komplett richtig zu erkennen. Von der hohen Redundanz wird gerade dann Gebrauch gemacht, wenn nur Teile des Satzes verstanden werden wie zum Beispiel bei Rauschbedingungen und Schwerhörigkeit. Somit würde man bei dieser Art der Sprichwörterveränderung die Verstehbarkeit tendenziell verschlechtern. Bei Verwendung von unveränderten Sprichwörtern und Redewendungen wird die Verstehbarkeit tendenziell verbessert. Da sie Alltagssprache repräsentieren, sollten derartige Sätze bei einem Sprachverständnistest vertreten sein, der den Kriterien der Alltagstauglichkeit gerecht werden will. So wurde in der Vorarbeit von Angermüller (unveröff. Diss.) festgelegt, dass Sprichwörter in ihrer originalen Form in das Sprachtestmaterial aufgenommen werden sollten.

Bei der Betrachtung der Formelhaftigkeit zeigte sich ein hochsignifikanter Zusammenhang bzgl. des Verständnismertes (s. Abbildung 3.13 auf S. 29). Es ist anzumerken, dass die Einteilung der Sätze nach ihrem Grad der Formelhaftigkeit im Rahmen dieser Arbeit durchgeführt wurde. Dementsprechend könnte dies auf den ersten Blick eine subjektive Einteilung sein. Da der Autor aber selbst den Kriterien entspricht, die für die teilnehmenden Probanden aufgestellt wurden, ist anzunehmen, dass diese subjektive Einteilung das Empfinden der Probanden größtenteils widerspiegelt. Allerdings muss man hier darauf hinweisen, dass sich die Anzahl der Sätze in den jeweiligen Kategorien sehr stark unterscheidet. Darüber hinaus gilt: je kleiner die Anzahl an Meßdaten in einer Kategorie, desto kritischer ist eine statistische Aussage darüber zu betrachten. Weiterhin ist auf einen anderen Zusammenhang hinzuweisen: genau genommen müsste man den Einfluss der Formelhaftigkeit unabhängig von der Satzkategorie und der Satzlänge betrachten. Denn diese beiden Faktoren nehmen ebenfalls Einfluss auf die Verstehbarkeit. Da nun aber in den Formelhaftigkeit-Kategorien „mäßig“ bzw. „stark“ nur wenige Messdaten vorhanden waren, war diese Betrachtung nicht sinnvoll. Aufgrund der hochsignifikanten Unterschiede ($p = 0,000018$) kann aber davon ausgegangen werden, dass die Formelhaftigkeit Einfluss auf das Satzverständnis hat. Der Formelhaftigkeit wurde bei den Testergebnissen insofern Rechnung getragen, dass die Listen nach der Schwierigkeit der Sätze eingeteilt wurden; darüber hinaus wird die Schwierigkeit der Sätze indirekt auch durch die Formelhaftigkeit mitbestimmt.

4.4.6 Einfluss des Inhalts der Satzblöcke

Es ist festzuhalten, dass die Sätze anfänglich auf die Blöcke zufallsverteilt wurden, so dass mögliche Unterschiede zwischen den Blöcken von vornherein nicht ausgeschlossen waren. Dies war auch nicht das Ziel der Zufallsverteilung. Durch die Methodik der Darbietung (s. Abschnitt 2.3.2 auf S. 12) ist es weitgehend auszuschließen, dass externe Einflüsse zu diesen Unterschieden führten. Eine Möglichkeit der Beeinflussung liegt zum einen in der zufälligen Verteilung der schwierigeren und leichteren Sätze in bestimmte Blöcke und zum anderen in der individuellen Streuung der Probanden begründet. Da jeder Proband nur ein Drittel der Gesamtmenge der Sätze abgehört hat, kann aber ein Vergleich zwischen den Probanden nicht genauer durchgeführt werden.

4.5 Satzlautstärke und deren Einfluss

4.5.1 Beobachtung

Beim Abhören der Sätze im Rauschen fiel auf, dass einige Sätze stärker veräuscht erschienen als andere und folglich schlechter verstehbar waren. Dieser subjektive Eindruck wurde sowohl vom Autor als auch von einigen Probanden geäußert. Deshalb erscheint es sinnvoll diesen Aspekt genauer zu betrachten.

Das Rauschen wurde durch ein von Dipl.-Ing. Brill geschriebenes Matlab-Skript den Sätzen hinzugefügt, welches folgendermaßen arbeitete: Als erstes wurden alle Sätze auf dieselbe zuvor ermittelte Lautstärke gerechnet. Als Referenzwert für die Berechnung galt für jeden Satz der höchste SPL-Wert im Zeitverlauf. Das heißt, Sätze, deren maximaler SPL hoch war, wurden leiser gerechnet und Sätze, deren Maximalpegel eher niedrig war, wurden lauter gerechnet. Das bedeutete, dass der Maximalpegel über die Lautstärke des ganzen Satzes entschied. Nach Abschluss der Berechnung waren alle Maximalpegel genau gleich laut und zwar 65 dB. Danach wurden alle Sätze zusätzlich mit einem Rauschen von -10 dB belegt, so dass der Maximalpegel nun eine Lautstärke von 65 dB plus 10 dB Rauschen aufwies. Alle anderen niedrigeren SPL-Pegel desselben Satzes wurden so mit Rauschen belegt, dass die Gesamtlautstärke über den ganzen Zeitverlauf 75 dB betrug. Im Prinzip entspricht diese Methodik zwei von einander unabhängigen Lautsprecherwiedergaben: dem konstanten Rauschen von 75 dB und dem darin dargebotenen Satz mit einem SPL-Spitzenwert

von 65 dB. Dies führte folglich dazu, dass Sätze, die eine ausgeprägte SPL-Spitze im Vergleich zum restlichen SPL-Verlauf aufwiesen, außerhalb dieser Spitze stärker verrauscht wurden, während Sätze, die einen relativ gleichmäßigen SPL-Verlauf zeigten, relativ gleichmäßig und daher weniger stark verrauscht waren.

Diese Tatsache ließ vermuten, dass ein Zusammenhang zwischen dem SPL-Spitzenwert und der Verstehbarkeit des Satzes vorlag. Deshalb schien es sinnvoll, diesen Sachverhalt genauer zu betrachten. Dazu wurden die Sätze im Bereich über 50 dB untersucht. Es wurde aufgelistet, wie oft welcher Pegel im Satzverlauf angenommen wurde. Nun wurde die Differenz zwischen Maximalpegel und dem am häufigsten vorkommenden Pegel, der als repräsentativer SPL des restlichen Satzes betrachtet werden konnte, gebildet. Dabei traten Differenzunterschiede von bis zu 20 dB auf. Die SPL-Differenz jedes Satzes wurde mit dem jeweiligen Verständniswert in Bezug gesetzt, um einen möglichen Zusammenhang überprüfen zu können. Graphisch ist dies in Abbildung 4.1 auf S. 42 nachzuvollziehen. Jeder Punkt repräsentiert dabei einen Satz. Der Rangkorrelationskoeffizient (Kendalls Tau) errechnete sich zu $-0,282508$. Das bedeutet, dass mit zunehmender Differenz die Verstehbarkeit des Satzes abnahm. Diese Korrelation ist mit $p < 1,0 \times 10^{-6}$ hochsignifikant. Das graphische Ergebnis einer durchgeführten linearen Regression wird durch die Gerade in Abbildung 4.1 auf S. 42 dargestellt.

Die Ergebnisse zeigen, dass ein Zusammenhang zwischen dem Zeitverlauf des SPLs und der Satzverständlichkeit besteht.

4.5.2 Einfluss durch Aufsprache

Ein Grund für die Unregelmäßigkeit des SPL-Verlaufs könnte in der Aufsprache selbst liegen. Im Gegensatz zu dieser Versuchsreihe wurden die originalen HSM-Sätze (Hochmair-Desoyer et al., 1997) von einem professionellen Sprecher aufgesprochen. Bei dem vorliegenden Test wurde die Aufsprache durch einen nicht professionellen Sprecher – Dipl.-Ing. Brill – durchgeführt. Die genauen Bedingungen für die hier verwendete Aufsprache sind in der Dissertation von Angermüller (unveröff. Diss.) nachzulesen. Hier sei kurz darauf hingewiesen, dass die Aufsprachen unter standardisierten Bedingungen mit höchster Sorgfalt durchgeführt und bearbeitet wurden. Ein professioneller Sprecher ist sicher in der Lage, eine sehr gut artikulierte Aussprache zu gewährleisten, so dass

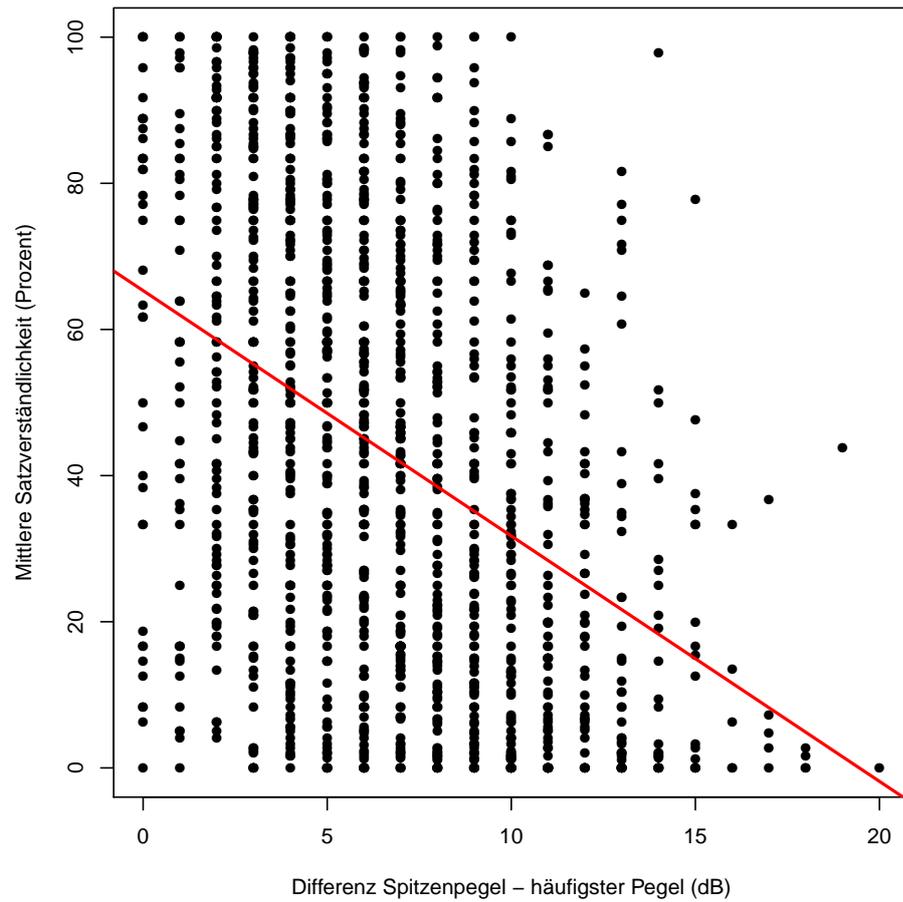


Abbildung 4.1: Korrelation von SPL-Differenz und Verständniswert

die Worte und Sätze gut verstanden werden. Ein nicht professioneller Sprecher kann dieses Niveau nur schwer erreichen, so dass allein durch die Erfahrung des Sprechers Unterschiede zustande kommen können. So können durch verschiedene Sprecher Unterschiede in der Verständlichkeit von 10-20 % (Martin, 1997) auftreten. Dies wird hauptsächlich auf geschlechtsspezifische Unterschiede auf Sprecherebene, aber auch auf regionale Unterschiede wie Akzente und Dialekte zurückgeführt. Geschlechtsspezifische Unterschiede lassen sich hier nicht finden, allerdings spricht Lehnhardt (2009) im Zusammenhang mit dem originalen HSM-Test von „Hochdeutsch mit leicht bayrischer Färbung“. Weiterhin berichtet Martin (1997), dass selbst ein einziger Sprecher nicht gleichmäßig artikulieren kann, so dass allein dadurch Unterschiede von bis zu 10 % zustande kommen können.

Zudem besitzt ein professioneller Sprecher wahrscheinlich eine homogenere Sprache, so dass daraus regelmäßigeren SPL-Verläufe resultieren können. Bei den hier verwendeten Sätzen wurde eine Schwankungsbreite der SPL-Verläufe von bis zu 20 dB im Bereich über 50 dB SPL gefunden. Wie gezeigt, nimmt das Satzverständnis mit zunehmender SPL-Differenz ab. In gewisser Weise liegen diese SPL-Unterschiede in der Natur der Spracheigenschaften. Niemeyer und Beckmann (1962) äußern sich so, dass die Sprachlautstärke eines Satztests kein konstantes Niveau haben kann, weil die Dynamik ein integrales Merkmal der lebendigen Sprache darstellt. Trotzdem wird von Niemeyer (1967) für die Brauchbarkeit sprachaudiometrischer Tests eine gleichmäßige Lautstärke gefordert. So gelang es dem Rundfunk-Sprecher, der den Marburger Satztest aufsprach, dass betonte Silben von Satz zu Satz eine maximale Differenz von 5 dB aufwiesen. Auch Lehnhardt (2009) fordert innerhalb einer Testgruppe denselben mittleren konstanten Schallpegel. Denn er stellte fest, dass bei lauter Stimme die Vokale in ihrer Lautstärke die Konsonanten überwogen und bei leiser Stimme die Konsonanten die Vokale. Deshalb wird von ihm ein gleichmäßiger mittlerer konstanter Schallpegel gefordert.

Interessant wäre es, in Zukunft die originalen HSM-Sätze und die gleichen neu aufgesprochenen Sätze - jeweils ohne Störgeräusch - einem Probanden darzubieten und ihn zu bitten, alle Sätze paarweise auf gleiche subjektive Lautheit einzustellen. So könnte man die resultierenden maximalen SPL-Spitzenwerte, wie auch die SPL-Differenzen beider Aufsprachen vergleichen und eventuell Rückschlüsse auf deren Einflüsse auf das Sprachverstehen ziehen. Die jeweiligen Satzpaare könnte man im Folgenden mit dem Störgeräusch versehen und beob-

achten inwiefern sich deren Verständlichkeit verändert. Daraus ließen sich wiederum Rückschlüsse ziehen, wie stark welche Kriterien einer Aufsprache durch das Rauschen beeinflusst werden.

4.5.3 Einfluss durch Lautstärkenangleichung

Es gibt einen weiteren Aspekt, der möglicherweise Einfluss auf das Sprachverstehen insbesondere im Störlärm nimmt: Die Methodik der Lautstärkenangleichung. Diese wurde bei der vorliegenden Arbeit durch mathematische Berechnung – mit dem SPL-Spitzenwert als Referenz – durchgeführt. Bei dem originalen HSM-Satztest wurde die Lautstärkenangleichung aller Sätze durch einen Tontechniker rein subjektiv durchgeführt. Welche Faktoren des SPLs im Verlauf eine Rolle für die Lautstärkeempfindung spielen ist nicht bekannt. Im Rahmen dieser Arbeit wurde festgestellt, dass nicht allein der Maximalpegel für das Lautstärkeempfinden entscheidend ist. Denn dieser wurde bei der Satzdarbietung gleichgesetzt. Trotzdem wurden subjektiv empfunden unterschiedlich laut ver-rauschte Sätze als Ergebnis erzielt.

Es wäre sinnvoll, die nicht verrauschten Sätze, die mit der hier angewandten Methode gleich laut gerechnet wurden, Probanden unverrauscht darzubieten. Es ist zu erwarten, dass alle Sätze zu 100 % verstanden werden, wenn ein Spitzenpegel von 65 dB vorhanden ist. Denn bei einer maximalen Schwankungsbreite von 20 dB zwischen dem Maximalpegel und dem häufigsten SPL-Wert ist immer noch ein SPL von 45 dB zu erwarten. Um Unterschiede in der Lautstärkeempfindung festzustellen, könnte man die Probanden auffordern, die Sätze subjektiv gleich laut zu stellen. So könnte man beobachten, ob die verwendete Methode der Lautstärkenangleichung auch im unverrauschten Zustand zu unterschiedlichen Lautstärkeempfindungen führt. Weiterhin könnte untersucht werden, ob eine Korrelation zwischen der Lautstärkeempfindung ohne Störlärm und der Lautstärkeempfindung im Störlärm existiert.

4.5.4 Fazit

An dieser Stelle bleibt ungeklärt inwieweit die Art der Aufsprache und inwieweit die Art des Verrauschens Einfluss auf die Verstehbarkeit der Sätze nimmt. Hier treten beide Aspekte gekoppelt auf.

Es konnte gezeigt werden, dass zunehmende SPL-Differenzen zu schlechte-

rem Satzverständnis im Rauschen führten. Deshalb schließt sich der Autor den Forderungen von Lehnhardt (2009) und Niemeyer und Beckmann (1962) an, dass eine homogene Aufsprache die Grundlage für zukünftige Tests darstellen sollte.

Weiterhin gibt es Grund zur Annahme, dass der maximale SPL-Spitzenwert nicht allein für die Satzlautstärke verantwortlich ist, sondern komplexere Einflüsse eine Rolle spielen. In Zukunft könnten diese Einflüsse genauer untersucht werden, mit der Absicht eine Methodik der Lautstärkenberechnung zu entwickeln, die diesen Einflüssen gerecht wird und gleichzeitig objektiv und reproduzierbar ist.

4.6 Listenzusammenstellung

Eine weitere Aufgabe dieser Arbeit war es, aus den hier getesteten Sätzen Listen so zusammenzustellen, dass alle Listen genau gleich schwierig waren. Dabei konnten die Listen nicht willkürlich eingeteilt werden. Schon beim originalen HSM-Satztest (Hochmair-Desoyer et al., 1997) gab es gewisse Voraussetzungen, die hier übernommen wurden. So war vorgeschrieben, wieviele Sätze welcher Satzlänge in jeder Liste vorkommen durften und welcher Satzkategorien sie angehören mussten (s. Abschnitt 2.4 auf S. 15). Auch in der Arbeit von Angermüller (unveröff. Diss.), aus der das Sprachmaterial für diese Arbeit entnommen ist, wurde diese Einteilung vorausgesetzt.

In dieser Arbeit ist es gelungen unter den genannten Voraussetzungen eine Methodik zu entwickeln, die viele exakt gleich schwierige Listen lieferte und damit zielführend war. Wenn auch die Aufsprache bzw. das Verrauschen der Sätze zu nicht homogenen Bedingungen bzgl. der Satzverständniswerte führten, ist die Methodik der Listeneinteilung unabhängig davon zu betrachten. Denn sie könnte in Zukunft auf weitere Testversuche problemlos übertragen werden.

5 Zusammenfassung

In der Vergangenheit wurde des öfteren Kritik am HSM-Satztest dahingehend geäußert, dass durch Wiederholungseffekte und Inhomogenität der Testlisten die Objektivität des Tests eingeschränkt sei. Auf diese Kritik sollte mit einer Erweiterung des HSM-Satztestmaterials reagiert werden.

Dazu wurden 1388 alte und neue Sätze zusammengestellt und aufgesprochen (Angermüller (unveröff. Diss.)). In der vorliegenden Arbeit wurden diese Sätze evaluiert und eine Methode entwickelt, mit der die Sätze so auf Listen verteilt werden konnten, dass am Ende alle Listen den gleichen Schwierigkeitsgrad aufwiesen.

Um die Situation eines Hörbehinderten, für den derartige Tests vorgesehen sind, zu imitieren, wurden anfänglich alle Sätze auf gleiche Lautstärke kalibriert und mit einem Rauschen versehen. Anschließend wurden sie 36 normalhörenden Probanden dargeboten, um jedem Satz einen Verständniswert zuzuweisen. Jeder Proband hörte ein Drittel der Gesamtsatzmenge, so dass jeder Satz von jeweils 12 Probanden beurteilt wurde. Basierend auf den durchschnittlichen Satzverständniswerten wurde eine Methode entwickelt, die die Sätze den Listen nach bestimmten Vorgaben zuwies, so dass am Ende Listen mit exakt gleicher Schwierigkeit als Ergebnis resultierten.

Die Methode der Listeneinteilung erwies sich als so leistungsfähig, dass ein zusätzliches Kriterium berücksichtigt werden konnte. Auch die Streubreite der Verständniswerte innerhalb der Listen konnten einander sehr weitgehend angeglichen werden. Die Methode ist auch auf andere Sprachverständnis tests übertragbar, sofern einige Voraussetzungen erfüllt sind.

Es stellte sich allerdings heraus, dass die Sätze von vielen Probanden als unterschiedlich laut wahrgenommen wurden. Dies kann einerseits an der Aussprache selbst gelegen haben, zum anderen ist es möglich, dass die Methode der Lautstärkenkalibrierung dafür verantwortlich ist. Dies zu untersuchen und gegebenenfalls zu verbessern sollte Aufgabe von zukünftigen Untersuchungen sein.

Literaturverzeichnis

Angermüller I. *Neuentwurf und Aufsprache eines Satzverständnistests*. Dissertation, Universität Würzburg, unveröff. Diss.

Bangert H. 1980. Probleme der Ermittlung des Diskriminationsverlustes nach dem Freiburger Sprachtest. *Audiologische Akustik*, 19:166–170.

Brill S. Persönliche Mitteilung, 2008.

Domes C. *Evaluierung und Weiterentwicklung eines Prosodieverständnistests*. Dissertation, Universität Würzburg, Am Hubland, 97074 Würzburg, 2009. URL <http://www.opus-bayern.de/uni-wuerzburg/volltexte/2010/4901>. [Stand: 02.08.2010].

Dunn C., Tyler R. S., Oakley S., Gantz B. J., und Noble W. Jun 2008. Comparison of speech recognition and localization performance in bilateral and unilateral cochlear implant users matched on duration of deafness and age at implantation. *Ear and Hearing*, 29(3):352–359. URL <http://dx.doi.org/10.1097/AUD.0b013e318167b870>. [Stand: 02.08.2010].

Grebe H. P. *Untersuchungen mit dem HSM-Satztest zum Sprachverständnis im Lärm bei Normalhörenden um 50 Jahre*. Dissertation, Universität Würzburg, 2000.

Hahlbrock K.-H. *Grundlagen und praktische Anwendungen einer Sprachaudiometrie für das deutsche Sprachgebiet*. Thieme-Verlag, Stuttgart, 1957.

Hellbrück J., M.-C.Oguey , und Seiler C. 1984. Sind geschlechtsspezifische Unterschiede in der Lautstärkenempfindung ein Artefakt der Gehörganggröße? *Zeitschrift für experimentelle und angewandte Psychologie*, Band 31, Heft 3: 439–446.

Helms J. und Müller J. Jan 1999. Die Auswahl eines Cochlea-Implantats und die Ergebnisse der Implantation. *Laryngorhinootologie*, 78(1):12–13. URL <http://dx.doi.org/10.1055/s-2007-996820>. [Stand: 02.08.2010].

- Hirschfelder A., Gräbel S., und Olze H. Mar 2008. The impact of cochlear implantation on quality of life: the role of audiologic performance and variables. *Otolaryngology - Head and Neck Surgery*, 138(3):357–362. URL <http://dx.doi.org/10.1016/j.otohns.2007.10.019>. [Stand: 21.07.2010].
- Hochmair-Desoyer I., Schulz E., Moser L., und Schmidt M. Nov 1997. The HSM sentence test as a tool for evaluating the speech understanding in noise of cochlear implant users. *American Journal of Otology*, 18(6 Suppl):S83.
- Hutchinson K. M. 1989. Influence of sentence context on speech perception in young and older adults. *Journal of Gerontology: Psychological Science*, 44 No 2:36–44.
- International Organization for Standardization . Acoustics – Statistical distribution of hearing thresholds as a function of age, 2000. URL http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=26314. [Stand: 31.03.2011].
- International Telecommunication Union . ITU-T Recommendation G227; Conventional Telephone Signal. Blue Book, Fascicle III.2, 11 1988.
- Kollmeier B., Müller C., Wesselkamp M., und Kliem K. Weiterentwicklung des Reimtests nach Sotscheck. In *Moderne Verfahren der Sprachaudiometrie, Band 1*, Seiten 216–236. Median Verlag von Killisch-Horn GmbH, Heidelberg, 1992.
- Kollmeier B. und Wesselkamp M. 1997. Development and evaluation of the german sentence test for objective und subjective speech intelligibility assessment. *Journal of the Acoustical Society of America (JASA)*, 102:2412–2421.
- Laszig R., Aschendorff A., Stecker M., Müller-Deile J., Maune S., Dillier N., Weber B., Hey M., Begall K., Lenarz T., Battmer R.-D., Böhm M., Steffens T., Strutz J., Linder T., Probst R., Allum J., Westhofen M., und Doering W. Nov 2004. Benefits of bilateral electrical stimulation with the nucleus cochlear implant in adults: 6-month postoperative results. *Otology & Neurotology*, 25(6): 958–968.
- Lehnhardt E. Sprachaudiometrie. In *Praxis der Audiometrie*, Kap. 13: Sprachaudiometrie, Seiten 147–164. Georg-Thieme-Verlag, Stuttgart, 9. Ausgabe, 2009.

- Lehnhardt E. und Laszig R., Hrsg. *Praxis der Audiometrie*. Georg-Thieme-Verlag, Stuttgart, 2009.
- Lindgren R. und Lindblom B. 1983. Speech perception processing. *Scandinavian Audiology*, 18:57–70.
- Martin M. *Speech Audiometry*, Kap. 2: Towards a theory of speech audiometry tests, Seiten 34–62. Whurr Verlag, London, 1997.
- Morton J. 1964. A preliminary functional model for language behaviour. *International Journal of Audiology*, 3:216–225.
- Mrowinski D. und Scholz G. *Audiometrie: eine Anleitung für die praktische Hörprüfung*. Thieme-Verlag, Stuttgart, 3. Ausgabe, 2006.
- Müller J., Schön F., und Helms J. 2002. Speech understanding in quiet and noise in bilateral users of the MED-EL COMBI 40/40 cochlear implant system. *Ear and Hearing*, 23:198–206.
- Müller-Deile J. 2009. Sprachverständlichkeitsuntersuchungen bei Kochleaimplantatpatienten. *HNO*, 57(6):580–591. ISSN 0017-6192. [Stand: 02.08.2010].
- Niemeyer W. 1967. Sprachaudiometrie mit Sätzen. *HNO*, 15:335–343.
- Niemeyer W. und Beckmann G. 1962. Ein Sprachaudiometrischer Satztest. *Archiv Ohr Nase Kehlkopfheilkunde*, 180:742–49.
- Oh S.-H., Kim C.-S., Kang E. J., Lee D. S., Lee H. J., Chang S. O, Ahn S.-H., Hwang C. H., Park H. J., und Koo J. W. 2003. Speech perception after cochlear implantation over a 4-year time period. *Acta Oto-laryngologica*, 123 (2):148–153.
- Sachs L. und Hedderich J. *Angewandte Statistik: Methodensammlung mit R*. Springer-Verlag, Berlin Heidelberg, 7. Ausgabe, 2009.
- Schlöndorff G. und Tegtmeier W. 1971. Ein audiometrischer Test zur Bestimmung des zentralen Aufwandes beim Sprachverstehen. *Laryngo-Rhino-Otologie*, 50:663–667.
- Schultz-Coulon H.J. 1973. Über die Bedeutung des Umweltgeräusches für den Hochtonschwerhörigen. *HNO*, 21:26–32.

- Schultz-Coulon H.J. 1974. Sprachaudiometrie mit Sätzen und Geräusch. *Laryngo-Rhinologie*, 53:734–48.
- Sotschek J. 1982. Ein Reimtest für Verständlichkeitsmessungen mit deutscher Sprache als ein verbessertes Verfahren zur Bestimmung der Sprachübertragungsrate. *Fernmelde-Ingenieur*, 36:1–84.
- Strohmaier C. J. *Sprachverstehen im Störlärm mit dem Innsbrucker Satztest in der HSM-Edition auf Compact-Disc bei Normalhörenden*. Dissertation, Universität Würzburg, 1998.
- Tschopp K. und Ingold L. Die Entwicklung einer deutschen Version des SPIN-Tests (Speech Perception in Noise). In *Moderne Verfahren der Sprachaudiometrie, Band 1*, Seiten 311–329. Median Verlag von Killisch-Horn GmbH, Heidelberg, 1992.
- Wedel von H. 1986. Entsprechen die sprachaudiometrischen Untersuchungsverfahren den heutigen Anforderungen in Klinik und Praxis? *HNO*, 34:71–74.
- Wagener K., Brand T., und Kollmeier B. 1999. Entwicklung und Evaluation eines Satztests in deutscher Sprache III: Evaluation des Oldenburger Satztests. *Zeitschrift für Audiologie*, 38:86–95.
- Wesselkamp M., Kliem K., und Kollmeier B. Erstellung eines optimierten Satztestes in deutscher Sprache. In *Moderne Verfahren der Sprachaudiometrie, Band 1*, Seiten 330–343. Median Verlag von Killisch-Horn GmbH, Heidelberg, 1992.
- Zimmermann M. *Verstehen im Störlärm mit HSM-Satztest bei Hörgeräteträgern zum Vergleich der Satzgruppen*. Dissertation, Universität Würzburg, 2000.

Danksagung

Mein herzlicher Dank gilt zunächst allen Probanden, die bereitwillig Zeit und Mühe aufgewandt haben, um bei der Versuchsdurchführung aktiv mitzuwirken.

Weiterhin danke ich Herrn Prof. Dr. med. Dr. h. c. Rudolf Hagen, Direktor der Klinik und Poliklinik für Hals-, Nasen- und Ohrenkrankheiten, plastische und ästhetische Operationen, für die Annahme als Doktorand sowie für die Möglichkeit die Klinikeinrichtung zur Versuchsdurchführung nutzen zu können.

Mein Dank für die freundliche Übernahme des Referats und Korreferats gilt Herrn Prof. Dr. med. Dr. h. c. Rudolf Hagen und Herrn PD. Dr. Dipl.-Ing. Mario Cebulla.

Für die Vergabe des Themas sowie für die umfangreiche und zeitintensive Betreuung danke ich Herrn Dipl.-Ing. Stefan Brill sehr. Mit seinem Wissen und seiner Erfahrung, seiner Gewissenhaftigkeit und Präzision hat er mich in das wissenschaftliche Arbeiten eingeführt und bei meiner Dissertation erfolgreich begleitet. Stets war er bereit offene Fragen zu beantworten und Ideen zu diskutieren.

Außerdem ist es mir ein Anliegen meinen Eltern für die Zeit und Unterstützung zu danken, die sie mir für diese Arbeit gewährt haben.

Lebenslauf

Name: Nicolai Lenz
Geboren am: 2. November 1981
In: Lörrach
Konfession: evangelisch
Familienstand: ledig

Schulbildung:

1988 - 1992 Grundschule: Hellbergschule
1992 - 2001 Hebel-Gymnasium Lörrach
06/2001 Abitur

Zivildienst:

09/2001 - 06/2002 Helen-Keller-Schule Maulburg, Schule für Körper- und Geistigbehinderte

Studium:

10/2002 - 08/2002 Magisterstudium, Universität Konstanz
Hauptfach: Französische Literaturwissenschaft
Nebenfächer: Volkswirtschaftslehre/Politikwissenschaft
10/2003 - 03/2004 Architekturstudium, Universität Weimar
04/2004 - 06/2010 Studium der Humanmedizin, Universität Würzburg
03/2006 Physikum
06/2010 Staatsexamen
06/2010 - 01/2011 Promotionsstudent

Berufsausübung:

seit 02/2011 Assistenzarzt