

**Aus der Klinik und Poliklinik für Hals- Nasen- und Ohrenkrankheiten,
plastische und ästhetische Operationen
der Universität Würzburg
Direktor: Professor Dr. med. Rudolf Hagen**

**Griechischer Satztest nach dem Vorbild
des HSM-Tests**

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

**Eleni Stavrianou
aus
Athen**

Würzburg, Juni 2010

Referent: Prof. Dr. med. Rudolf Hagen
Koreferent: Priv.-Doz. Dr. Ing. habil. Mario Cebulla
Dekan: Prof. Dr. med. Matthias Frosch

Tag der mündlichen Prüfung: 29.11.2010

Die Promovendin ist Ärztin.

Meiner Familie

Inhaltsverzeichnis

| | | |
|------|---|----|
| 1 | Einleitung | 1 |
| 2 | Aspekte der Sprachaudiometrie | 4 |
| 2.1 | Einführende Bemerkungen | 5 |
| 2.2 | Entwicklung der Sprachaudiometrie | 5 |
| 2.3 | Diskutierte Nachteile von Sprachtests | 17 |
| 2.4 | Rauschen | 18 |
| 2.5 | Anforderungen an den Aufbau eines Satztests | 20 |
| 2.6 | HSM-Satztest | 21 |
| 2.7 | In Griechenland gebräuchliche Sprachtests | 22 |
| 2.8 | Ziel der Arbeit | 23 |
| 3 | Vorbereitende Arbeiten | 24 |
| 3.1 | Materialsammlung | 24 |
| 3.2 | Generierung des Rauschsignals | 25 |
| 4 | Vorstudie | 26 |
| 4.1 | Material und Methoden | 26 |
| 4.2 | Versuchsdurchführung | 27 |
| 4.3 | Ergebnisse | 29 |
| 5 | Hauptstudie | 30 |
| 5.1 | Material, Methoden, Versuchsdurchführung | 30 |
| 5.2 | Ergebnisse | 36 |
| 6 | Konstruktion der endgültigen Listen | 43 |
| 7 | Diskussion | 47 |
| 8 | Zusammenfassung | 56 |
| 9 | Literaturverzeichnis | 57 |
| 10 | Tabellarischer Anhang | 69 |
| 10.1 | Verständlichkeit der Listen | 69 |
| 10.2 | Satzlisten | 70 |

1 Einleitung

Auge und Ohr eröffnen uns einen großartigen und tief erlebten Zugang zu dieser, unserer Welt. Zu erblinden oder zu ertauben empfinden die Betroffenen deshalb als einen besonders schmerzlichen Verlust. Obwohl sich die meisten Menschen als Augenwesen sehen, zeigt ein Vergleich zwischen blind und taub geborenen Kindern, dass Kinder, die nicht hören können, wesentlich mehr in ihrer Entwicklung beeinträchtigt werden als blinde. Von Geburt an blinde Kinder vermögen ihr Gehör derart zu schärfen, dass das Fehlen der visuellen Wahrnehmung teilweise kompensiert werden kann. Anhand reflektierter Schalle können Gegenstände erkannt und ihre Lage im Raum erhört werden. Taub geborene Kinder waren in der Vergangenheit sehr viel schlechter gestellt. Die nächsten Bezugspersonen folgten lange Zeit angeborenem Verhalten und behandelten ihr taubes Kind wie ein normal hörendes, indem sie versuchten, lautsprachlich mit ihm zu kommunizieren. Diese nicht angepasste Strategie führte dazu, dass die Kinder wegen der mangelnden Kommunikation in ihrer geistigen Entwicklung weit hinter ihren normalhörenden Altersgenossen zurückblieben. Wenig mitfühlend fand dieser Mangel auch im deutschen Sprachgebrauch seinen Ausdruck: Das Wort ‚doof‘, die niederdeutsche Entsprechung des hochdeutschen Wortes taub, wird heute in dem abwertenden Sinn von dumm gebraucht (Duden 1977).

Auch ein Erwachsener spürt, welche große Einbuße an Lebensqualität mit dem Verlust des Gehörs einhergeht. Wie bei taub geborenen Kindern leidet die zwischenmenschliche Kommunikation, wenn gesprochene Sprache nicht mehr verstanden werden kann. Nicht selten führt dies in die soziale Isolation.

Evolutionär gesehen ist die Anpassung unseres Gehörs an die Erfordernisse lautsprachlicher Kommunikation auf dem erreichten hohen Niveau eine einmalige und sehr junge Entwicklung. Dagegen können beinahe alle Lebewesen an Vibrationen des Untergrundes oder der Luft erkennen, wenn irgendwelche Gefahren drohen. Auch in dieser Funktion ist das Gehör für den Menschen ein überaus hilfreicher Sinn geblieben. Selbst im Schlaf, wenn unsere Augen geschlossen sind, bleibt das Ohr wachsam. Es scheint, dass es auf gewisse Geräusche sogar sehr empfindlich programmiert werden kann: Ein bekanntes

Beispiel hierfür liefern Mütter Neugeborener, welche bereits durch ein leises Wimmern des Kindes geweckt werden, während der Vater nebenan seelenruhig weiterschläft. In der Funktion des Warnens ist das Gehör dem Sehen in einem Punkt überlegen. Man hört, was um einen herum geschieht und nicht nur, was sich vor einem im Blickfeld abspielt. In diesem Zusammenhang war das Gehör dem Menschen auch in seiner Rolle als Jäger nützlich. Beute, die sich im dichten Unterholz oder in hohen Baumkronen verborgen hatte, konnte sich durch Geräusche verraten und aufgespürt werden.

Dieser kurze Abriss zeigt deutlich, wie wichtig es für den Menschen ist, hören zu können und wie schmerzlich der Verlust des Gehörs erlebt wird. Leider führt unsere moderne Lebensweise dazu, dass immer mehr Menschen unter Hörschäden leiden. Unsere Umwelt ist im Vergleich zu früher sehr viel lauter geworden. Erwiesenermaßen schadet dieser Lärm dem Ohr. Andererseits werden Dank der modernen Medizin die Menschen immer älter, und mit dem Alter lässt, von Ausnahmen abgesehen, auch das Gehör nach. Außerdem wächst auch mit dem Alter die Wahrscheinlichkeit zu erkranken, wovon auch das Gehör nicht verschont bleibt.

Was kann man tun, wenn jemandem ein solches Unbill zustößt? Zunächst gilt es, das Ausmaß der Hörstörung und die Ursachen zu ergründen. Sinnvollerweise kann sich Hilfe für die betroffenen Personen nur auf eine solide Diagnostik gründen. Soweit es um die Frage der Hörfähigkeit geht, kann man auf die unterschiedlichen Methoden der Audiometrie zurückgreifen. Gelegentlich, wie zum Beispiel bei der Versorgung mit Hörgeräten oder Cochlea-Implantaten, kann die Audiometrie auch Hilfen anbieten, um die Geräte optimal anzupassen oder einzustellen. Hierbei interessiert vor allem, wie gut Sprache unter verschiedenen Umweltbedingungen verstanden werden kann. Um hierauf eine quantitative Antwort geben zu können, bedient man sich unterschiedlicher sogenannter „Sprachtests“.

Für die Untersuchung von deutschsprachigen, mit Cochlea-Implantaten versorgten Patienten wurde der HSM-Satztest (Hochmair-Desoyer et al. 1997) entwickelt. Da zu Beginn dieser Arbeit in Griechenland ein dem HSM-Satztest entsprechender Test fehlte, aber an mehreren Kliniken zunehmend mehr

Patienten mit Cochlea-Implantaten versorgt werden, schien es angebracht, auch für das Griechische einen solchen Test zu entwickeln. Daraus leitet sich die Zielsetzung dieser Arbeit ab, nämlich die aufgezeigte Lücke zu schließen und einen evaluierten Sprachtest nach dem Vorbild des deutschen HSM-Satztests für das Griechische zu realisieren.

2 Aspekte der Sprachaudiometrie

Feldmann (2004) weist auf einen wichtigen Punkt für das Verständnis der historischen Entwicklung der Audiometrie hin. Solange keine Möglichkeit besteht, einem Tauben oder Schwerhörigen zu helfen, gleichzeitig aber das Handicap für jedermann ersichtlich ist, besteht kaum ein Grund, den Grad der Schwerhörigkeit quantifizieren zu wollen. Besteht aber ein Interesse, das Ausmaß einer Schwerhörigkeit zu quantifizieren, so gelingt dies nur, wenn mehrere Bedingungen erfüllt sind. Da quantitative Ergebnisse angestrebt werden, ist im weitesten Sinne eine Messung durchzuführen. Offensichtlich setzten die vorhandenen und einsetzbaren technischen Möglichkeiten hierfür Grenzen. Ein Blick in die Geschichte der Audiologie zeigt, dass die Technik allzeit wie ein beengendes Korsett wirkte. Aber Technik alleine reicht nicht aus, das gestellte Problem adäquat zu lösen. Bevor man eine Messung angehen kann, muss klar sein, was gemessen werden soll. Auf diese Frage scheint es, da ein komplexes System in einer komplexen Umwelt untersucht werden soll, bis zum heutigen Tag keine befriedigende Antwort zu geben. Schon allein die richtige Frage zu stellen, kann nur gelingen, wenn das dafür notwendige Wissen vorhanden ist. Notwendigerweise erscheint so die Entwicklung der Audiometrie als ein iterativer Prozess.

Ein wesentlicher Schritt hin zu größerer Reproduzier- und Vergleichbarkeit der Ergebnisse fand nach dem Zweiten Weltkrieg statt, da nunmehr die notwendigen technischen Hilfsmittel für audiometrische Untersuchungen zur Verfügung standen (Hahlbrock 1953, Papafragou 1996). Gleichzeitig bestand wegen der neu entwickelten gehörverbesserten Operationstechniken ein verstärktes Interesse an diesbezüglichen Daten. Andererseits stießen auch die Nachrichtentechniker in ihrem Arbeitsgebiet auf ähnliche Fragestellungen wie die Mediziner. Wesentliche Impulse zur Entwicklung der Sprachaudiometrie stammen daher aus diesem Technikbereich. Beispielsweise interessierte im Fernmeldewesen, wie gut die mit technischen Hilfsmitteln übertragene Sprache verstanden werden kann. Folgerichtig entwickelten die Techniker entsprechende Tests, um darauf eine Antwort geben zu können (Fletcher 1953, Egan 1948, Sotscheck 1976). In der Medizin interessieren zwar nicht die Eigenschaften eines technischen Gerätes, aber in der Frage, wie sehr ein Hör-

schaden die Verständlichkeit von sprachlichen Äußerungen beeinträchtigt, berühren sich die beiden Gebiete. Darüber hinausgehend möchte man natürlich gern aus den gewonnenen Informationen auf die Ursache, die Erkrankung selbst, schließen können. Auf beide Fragestellungen versucht die Audiometrie, eine Antwort zu geben.

2.1 Einführende Bemerkungen

Die Wurzeln der modernen Sprachaudiometrie reichen weit verzweigt tief in die Vergangenheit zurück. Bezold (1896) hat schon am Ende des 19. Jahrhunderts die Bedeutung der Sprache für diagnostische Zwecke herausgestellt. Er meinte, dass „die Sprache eine so vollkommene Zusammenstellung aller möglichen Lautkomplexe enthält, dass wir sie für unsere Hörproben erfinden müssten, wenn wir sie nicht schon hätten“ und „als regulären und zu einer raschen Übersicht führenden Hörmesser besitzen wir keinen vollkommeneren als die Sprache“. Mit Sprache als Reiz nutzt man die gleichen Lautkombinationen, die in gewohnter Weise auch im Alltag zum Zwecke der Kommunikation benutzt werden. Die Sprachaudiometrie stellt deshalb ein für den Prüfling leicht zu verstehendes Untersuchungsverfahren dar und erfüllt damit ein für die routinemäßige Anwendung wichtiges Kriterium (Bezold 1896). Lyregaard (1982) sieht in den Sprachtests ein für den Untersuchungszweck besonders adäquates Instrument, da das menschliche auditorische System insbesondere auf das Verstehen von Sprache ausgerichtet sei.

2.2 Entwicklung der Sprachaudiometrie

Die Darstellungen von Feldmann (1960, 2004), Hahlbrock (1970) und Bosman (1992) zeigen, dass die Anfänge der Sprachaudiometrie in einer Zeit weit vor Bezold liegen. Danach hat Pflingsten (1804) bereits vor etwa 200 Jahren in Kiel Sprache benutzt, um schwerhörige Patienten zu untersuchen. Damit ist die Sprachaudiometrie älter als die Tonaudiometrie. Pflingsten nutzte für seine Untersuchungen Eigenschaften der Vokale und Konsonanten. Das Ausmaß der Schwerhörigkeit versuchte er, durch eine Einteilung in drei Klassen zu beschreiben. Ähnlich verfährt wenig später Itard (1821) mit einer erweiterten Einteilung in fünf Klassen. Einen bedeutenden Schritt, den Hörverlust quantitativ sicherer

zu erfassen, taten Schmaltz (1846) und v. Tröltsch (1867), indem sie als Maß die sogenannte Hörweite benutzten und damit einen Bezug zu einer physikalisch messbaren Größe herstellten.

In seiner viel beachteten Arbeit von 1863 korreliert von Helmholtz den charakteristischen Klang eines Vokals mit Resonanzen im Mund-Rachen-Raum. Damit schafft er einen Zusammenhang zwischen dem Klang der Vokale und deren spektraler Energieverteilung. Diese Erkenntnis muss in einem Ohrenarzt zwangsläufig den Gedanken aufkommen lassen, die Hörbarkeit von Tönen und Geräuschen in unterschiedlichen Frequenzbereichen zu untersuchen. Damit war eine Entwicklung angestoßen, welche in die moderne Tonaudiometrie einmündete. Da die heute vorhandenen technischen Voraussetzungen für derartige Untersuchungen fehlten, glaubte man, in den Eigenschaften der Sprachlaute einen Zugang gefunden zu haben. So hat Wolf (1899) die von Helmholtzsche Analyse auf die Konsonanten übertragen, weil er überzeugt war, dass die Verständlichkeit der Sprache vor allem durch diese bestimmt wird. Er suchte nach Wörtern, die sich durch eine frequenzmäßig hohe oder tiefe Lage unterscheiden. Ihm kommt das Verdienst zu, durch seine Untersuchungen zur Hörweite in einer Allee auf die Tatsache hingewiesen zu haben, dass diese entscheidend von der Art des Stimulus abhängt. Die Holländer Zwaardemaker und Quix (1905) gingen noch einen Schritt weiter, indem sie versuchten, Wörter zu benutzen, die nicht nur gleiche Tonhöhe aufwiesen, sondern auch gleiche Lautstärke. Auch dieser Ansatz, dass sich aus der Verständlichkeit unterschiedlicher Laute nützliche Informationen extrahieren lassen, hat sich bis in die Moderne gehalten. Immer wenn Sprachtests nach phonetischen oder phonemischen Gesichtspunkten ausgewertet werden, zeigt sich eine neue Variante dieser alten Ideen.

Andererseits werden in einer Sprache auch nicht alle möglichen Lautkombinationen realisiert. So ist Reuter (1904), der für die deutsche Sprache Listen mit einsilbigen Wörtern erstellen wollte, auf sinnlose Wörter ausgewichen, weil er nicht genügend sinnvolle Wörter fand, die bezüglich der Intensität und Tonhöhe als äquivalent gelten konnten. Den Anstoß hierzu gaben seine Patienten. Obwohl er sie instruiert hatte, dass sie nur sinnvolle Wörter hören

würden, nannten sie immer wieder sinnlose Wörter. Sein Material bestand aus einsilbigen und zweisilbigen Wörtern, die sich jeweils nur durch einen „Wechsel-laut“ unterschieden.

Damit nimmt er einen Teil einer anderen Idee vorweg, die des Reimtests. Barany (1910) experimentierte zunächst auch mit Listen aus sinnlosen Wörtern und Silben, weil er glaubte, er könne auf diese Weise ausschließen, dass ein Testwort richtig erraten wird, obwohl es nur teilweise verstanden wurde. Er musste jedoch enttäuscht feststellen: „Selbst wenn man den Patienten wiederholt instruiert, er werde sinnlose Wörter hören, bemüht er sich immer wieder, sinnvolle Wörter zu verstehen, um so mehr, je weniger intelligent er ist“. Barany entschied sich daher letztlich wieder für einen Sprachtest mit sinnvollen Wörtern. Gleichzeitig wollte er zu leichter interpretierbaren Ergebnissen kommen. Die Lösung sah er darin, nur die Verständlichkeit eines einzigen Lautes zu prüfen. Entsprechend bestand sein Test aus Gruppen von einsilbigen und zweisilbigen Wörtern, wobei auch die zweisilbigen Wörter nur einen betonten Vokal besaßen. Die Wörter unterschieden sich nur in einem einzigen Laut, dem sogenannten „Wechsel-laut“, z. B. Wabe, Wade, Waffe, Ware, schaffe, Pfaffe. Falls der Wechsel-laut wie im gezeigten Beispiel ein Konsonant ist, so bezeichnet man heute einen solchen Test als Reimtest. Sein Test umfasste insgesamt 24 Wortgruppen.

Auf den ersten Blick mag es scheinen, als ob Reuter und Barany sich einander widersprechen würden. Dies ist nicht der Fall: Falls die Wahrnehmung eines Reizes auf irgendeine Art erschwert ist, beispielsweise durch eine Schwerhörigkeit, so kann ein sinnloses Wort also solches oder als sinnvolles verstanden werden, und alternativ kann die Antwort auf ein sinnvolles Wort ein sinnvolles oder sinnloses sein. Hören ist ein hochkomplexer Vorgang, und beides ist möglich.

Die Idee Baranys wurde von Lampert (1923) aufgegriffen, wobei er seinen Test nach den neuesten phonetischen Erkenntnissen, die Carl Stumpf 1926 erarbeitet hatte, zusammenstellte. Er schuf modern anmutende „Reimensembles“, siehe Sotscheck (1976a), deren „Mitglieder“ sich nur im Anlaut, Inlaut oder Auslaut unterschieden. Allerdings bewertete er jedes Testwort einzeln in

Bezug auf seine frequenzspezifische Besonderheit. Lambert folgte in dieser Hinsicht Wolf, der behauptet hatte, dass sich die Prüfung mit Sprache mehr auf die Sprachelemente als auf die lebendige Sprache selbst gründen sollte. Die hier aufscheinende Differenz berührt eine grundsätzliche Frage: Gilt auch für das Sprachverstehen, dass das Ganze mehr als die Summe seiner Teile ist? Oder kann das Ganze befriedigend durch die Funktion seiner Teile beschrieben werden?

Bezold erkannte, dass sich nicht alle Patienten auf dem gleichen sprachlichen Niveau bewegen und solche Unterschiede in der Sprachfertigkeit auf die Testergebnissen durchschlagen können (Bezold 1896). Die Testwörter sollten daher allen Patienten gleich bekannt und geläufig sein, damit sie mit gleicher Häufigkeit assoziiert werden. In den Zahlen glaubte er, eine solche Wortgruppe gefunden zu haben. Als Prüfmaterial schlug er daher die Verwendung der Zahlen von Eins bis Hundert vor. Er benutzt unterschiedliche Zahlen, um das Hörvermögen bei tiefen, mittleren und hohen Frequenzen beurteilen zu können. Trotz vieler Einwände wurde die Benutzung der zweistelligen Zahlwörter zur etablierten Praxis in der klassischen Hörweitenprüfung mit Flüster- und Umgangssprache (Feldmann 2004). Die Zahlwörter fanden auch Eingang in den Freiburger Sprachtest. In anderen Tests versuchte man, den Einfluss unterschiedlicher sprachlicher Fertigkeiten und Begabungen auf die Ergebnisse dadurch auszugleichen, dass man nur die in der jeweiligen Sprache am häufigsten benutzten Wörter in die Sprachtests aufnahm.

In den bisher beschriebenen Untersuchungsmethoden wurden die Testitems den Patienten vorgesprochen, und zwar mit unterschiedlichen Lautstärken. Daraus ergeben sich gleich mehrere Probleme. Flüstert man sehr leise, so verwischt sich der Unterschied zwischen Vokalen und Konsonanten. Bezold schlug daher vor, die Testwörter immer gleich laut vorzusprechen und die Lautstärke am Ohr des Prüflings dadurch zu ändern, dass man den Abstand zwischen Prüfer und Prüfling verändert. In der modernen Audiometrie verwirklicht man diesen Vorschlag, indem man eine Aufzeichnung des Sprachtests mit unterschiedlicher Lautstärke vorspielt.

Auf ein weiteres Problem hat nach Feldmann (2004) schon Kramer 1861 hingewiesen, nämlich dass es nicht immer gelingt, mit der gleichen Lautstärke und Modulation der Stimme zu sprechen. Fletcher (1953) benutzte daher bereits in den Zwanzigern des letzten Jahrhunderts bei den Bell Telephone Laboratories in den Vereinigten Staaten die nun vorhandene Möglichkeit, das Testmaterial auf elektromechanischem Wege zu präsentieren und auf diese Weise auch die Lautstärke zu kontrollieren und zu ändern. In Amerika erkannte man während des Zweiten Weltkrieges die Bedeutung einer guten Kommunikation für eine effiziente Kriegsführung. Die in diesem Zusammenhang entwickelten Methoden und Testmaterialien konnten auch für die Audiologie nutzbar gemacht werden. In diesem Zusammenhang entstand der Harvard-PB-Einsilbertest (phonetically balanced) von Egan (1948, 1957). Dieser umfasste 24 verschiedene Listen mit je 50 verschiedenen, sinnvollen, einsilbigen Wörtern. Die Häufigkeitsverteilung der Laute wurde der Verteilung im amerikanischen Englisch weitgehend angeglichen, um einen für die Sprache repräsentativen Test zu erhalten. 1948 publizierte Egan eine überarbeitete Version, welche aus 20 phonologisch angepassten Listen zu je 50 Einsilbern bestand (Egan 1948). Als sogenannte Harvard-PB-50-Monosyllabic Word Lists wurde dieser Test sowohl in der Nachrichtentechnik als auch in der Audiometrie angewandt. 1960 wurde er vom American National Standards Institut genormt (ANSI 1960).

Ebenfalls in Harvard wurde der sogenannte „Spondee-Test“, der aus Wörtern mit zwei gleich stark betonten Silben besteht, ausgearbeitet. Ausgehend von den Harvard-PB-50-Listen entwickelten Hirsh et al. (1952) einen Test speziell für die Audiometrie, den „Auditory Test W-22“. Vier phonologisch ausbalancierte Listen von je 50 einsilbigen Wörtern dienten als Ausgangsmaterial. Indem die Reihenfolge der Wörter in jeder Liste sechsmal vertauscht wurde, ergaben sich insgesamt 24 Listen zu je 50 Wörtern, von denen je sechs phonetisch äquivalent sind.

Einsilber einer speziellen Bauart bedienten sich Lehiste und Peterson in dem 1995 veröffentlichten CNC-Test (Lehiste 1959). Jedes Wort besteht aus drei Phonemen, einem Konsonanten im Anlaut, gefolgt von einem Vokal als Kern

(N) und einem Konsonanten im Auslaut. Auch sie haben bei der Auswahl der Wörter deren Häufigkeit im gesprochenen amerikanischen Englisch berücksichtigt, um der Abhängigkeit der Verständlichkeit vom Bekanntheitsgrad der Wörter Rechnung zu tragen.

Man hat versucht, die Tests aus dem angloamerikanischen Sprachraum ins Deutsche zu übertragen. Dies stieß jedoch auf Schwierigkeiten, weil sich die Sprachen nicht nur im Hinblick auf den Wortschatz unterscheiden. Wegen der Verschiedenheit der Sprachen ist es im Unterschied zur Tonaudiometrie bisher nicht gelungen, einen universellen Sprachtest zu entwickeln. Schubert scheint, wie Hahlbrock (1953) angibt, auf dem ersten Audiologenkongress in Stockholm internationale Sprachtests propagiert zu haben. Da man hierbei auf sinnlose Silben ausweichen müsste, womit, wie bereits erwähnt, Nachteile verbunden sind, stieß diese Anregung auf wenig Nachahmer. Inzwischen hat jede Sprachgemeinschaft ihre eigenen Tests entwickelt. Regionale Unterschiede, wie sie sich in den unterschiedlichen Dialekten manifestieren, können selbst schon im deutschen Sprachraum zu Schwierigkeiten führen.

Aufbauend auf dem Sprachtest von Hagerman (1982) für das Schwedische verfolgt Kollmeier eine interessante Alternative, gleichartige Tests für unterschiedliche Sprachen zu entwickeln. Der deutschen Version, dem sogenannten OLSA-Test (Wagener et al. 1999a, 1999b, 1999c, OLSA 2000) folgte inzwischen eine dänische Variante (Wagener 2003a, Wagener et al. 2003b). Die Tests bestehen aus strukturgleichen Fünfwortsätzen der Form: Vorname, Verb, Zahlwort, Adjektiv, Objekt. Bei der Synthese der Sätze wird für jede Wortart eines aus einem Vorrat von 10 Wörtern ausgesucht. Die so generierten Sätze sind zwar grammatikalisch richtig, aber die einzelnen Wörter stehen manchmal in einem ungewohnten Sinnzusammenhang.

Ursprünglich sollte mit diesem Testaufbau ein ganz anderes Problem gelöst werden. Es kommt immer wieder vor, dass bei umfangreichen Untersuchungen der Vorrat an Listen nicht ausreicht und Listen mehrfach vorgespielt werden müssen. Da Alltagssätze leicht zu behalten sind, können nur teilweise verstandene Sätze bei einer Wiederholung der Testlisten aus der Erinnerung heraus vervollständigt werden. Durch diesen Lerneffekt verliert der Test an Aussage-

kraft. Ein zweites Problem besteht darin, dass die Listen zwar für das bei der Entwicklung des Tests benutzte Kollektiv gleiche Verständlichkeit aufweisen, dies aber nicht für Probanden mit anderen Höreigenschaften gelten muss. Die Idee von Hagerman (1982) bestand darin, beide Probleme dadurch auszuräumen, dass erstens bei jedem Testdurchgang die gleichen Wörter benutzt werden, so dass bezüglich des verwendeten Wortmaterials zwischen zwei aufeinanderfolgenden Testdurchgängen kein Unterschied besteht, und zweitens wegen der zufälligen Anordnung der stets gleichen Wörter kein Lerneffekt zu erwarten ist. Untersuchungen (Hagerman, Kinnfors 1995) haben gezeigt, dass der Lerneffekt nicht ganz eliminiert werden konnte. Für die neueste Version dieser Testart, einen Test in Dänischer Sprache, gibt Wagener (2003a) für die ersten 8 Listen eine Verbesserung der Sprachverständlichkeitsschwelle von 2,2 dB an. Jedes Abhören fördert die Vertrautheit mit dem Wortinventar, und schließlich weiß man, falls man die Lautfolge (... efa ...) hört, die für sich keinen Sinn macht, dass es nur das verstümmelte Wort „Stefan“ sein kann. Da der Zuwachs an Verständlichkeit mit jeder Liste, die man abhört, geringer wird, lässt sich dieser Übungseffekt dadurch abschwächen, dass man der eigentlichen Messung mehrere Listen vorschaltet. In der audiologischen Routine sieht man davon ab, weil sich das Augenmerk auf den Probanden und nicht auf den Test richtet. Der größere Aufwand erscheint daher nicht sinnvoll und man verzichtet darauf.

Beim alltäglichen Sprechen beeinflussen sich Auslaut und Anlaut zweier aufeinander folgender Wörter gegenseitig. Diese Koartikulation wurde bei Hagerman noch nicht, in den nachfolgenden Tests jedoch berücksichtigt.

Die Ergebnisse eines Sprachtests können aber bereits von weniger auffälligen Begebenheiten beeinflusst werden, beispielsweise durch das Geschlecht des Sprechers oder durch die persönliche Note seiner Aussprache. Weitere Aspekte formulierte Schubert (1952) in einem Forderungskatalog: „1) Der Test muss die wichtigsten, in der deutschen Sprache vorkommenden Laute in derselben prozentualen Zusammensetzung zeigen, wie sie im Deutschen durchschnittlich auftreten. 2) Diese Forderung muss mit einem Minimum an Wörtern erzielt werden, um die Prüfdauer so kurz wie möglich zu halten. 3) Die Laute

sollen nicht in sinnlosen Silben, sondern in den am häufigsten vorkommenden Wörtern untergebracht sein, um möglichst wirklichkeitsgetreue, phonetische Verhältnisse zu erzielen und die Kombinationsgabe zu prüfen. 4) Die Anzahl der Laute in jeder Wortgruppe, die durch Pausen getrennt sind, sollen möglichst gleich sein. 5) Die Pause nach jeder Wortgruppe muss ausreichend lang sein, um dem Ohr Zeit zur Readaptation und Erholung von der Ermüdung zu geben. 6) Der Sprecher muss ein einwandfreies Hochdeutsch sprechen.“

Im Deutschen zählt der sogenannte „Freiburger Sprachtest“, entwickelt von Hahlbrock (1953, 1970), zu den Standards in der Audiologie. Amerikanischem Vorbild folgend entsprechen die Testwörter phonetisch der Häufigkeit in der deutschen Sprache. Diese Forderung ist nicht ganz trivial zu erfüllen. In jedem Fall können die Testlisten nur einen kleinen Ausschnitt aus dem Wortvorrat einer Sprache enthalten. Eine weitere Einschränkung ergibt sich daraus, dass nur geläufige Wörter in die Testlisten aufgenommen werden sollten. Im Deutschen stellen die einsilbigen Wörter nach der Auszählung von Meier die häufigste Wortart dar. Hahlbrock stützte seinen Test deshalb auf diese Wortart. Sein „Freiburger Sprachtest“ umfasst 20 Listen zu je 20 einsilbigen Wörtern. Da Einsilber wegen der geringen Redundanz schwer zu verstehen sind, stellte er dem Einsilbertest einen Test mit den leichter zu verstehenden Zahlen an die Seite. Dieser Test besteht aus 10 Listen zu je 10 zweistelligen Zahlen. Hahlbrock verwendete einen männlichen Sprecher, weil er für männliche Sprecher eine bessere Reproduzierbarkeit als für weibliche fand. Der Freiburger Sprachtest wurde basierend auf den Untersuchungen von Brinkmann (1974) zur DIN 45621 erhoben.

Offensichtlich lassen sich für einen Sprachtest viele Randbedingungen formulieren. Es überrascht daher nicht, dass immer wieder neue Schwerpunkte gesetzt wurden. Sotscheck (1976a, 1976b, 1985) haucht mit seinem Reimtest der Idee der Wechsellaudreihen neues Leben ein. Der Test hat gegenüber den im Fernmeldewesen ausgiebig benutzten Logatomtests den Vorteil, dass bei wiederholter Anwendung des Tests der Lerneffekt gering ausfällt. Ein anderer wesentlicher Vorteil besteht darin, dass sich der Test leicht automatisieren lässt. Wegen der vorgegebenen Alternativen können die Probanden ihre Antworten

mit Hilfe einer einfachen Zifferntastatur abgeben. Sotscheck bietet jeweils sechs sinnvolle, einsilbige Wörter, sogenannte Reimensembles, zur Auswahl an. Diese unterscheiden sich entweder im konsonantischen Anlaut, im Vokal des Inlauts oder im konsonantischen Auslaut. Im Anlaut oder Auslaut waren auch komplexe Lautgebilde, wie z. B. in dem Ensemble mach, Mast, Macht ..., zugelassen. Jede Liste umfasst 100 Wörter, davon bilden 34 den Anlautteil und je 33 den In- und Auslautteil. Leider induziert die Beschränkung auf sechs Antwortalternativen Varianz. Falls ein Wort falsch gehört wird, aber nicht in der Liste der Antworten vorkommt, so gibt es zwei Strategien, darauf zu reagieren. Zum einen kann man auf seinem Höreindruck beharren, woraus folgt, dass man keine Antwort abgibt. Das Wort wird dann als nicht verstanden gewertet. Auf der anderen Seite kann man aus der Tatsache, dass man das Wort nicht in der Liste findet, schließen, dass man sich verhöhrt hat. Naheliegenderweise wird man daraufhin nach einem ähnlich klingenden Wort suchen und dieses dann in der Liste markieren. Mit dem undeutlichen Höreindruck im Hinterkopf besteht eine nicht geringe Chance, die richtige Alternative zu treffen. Abhängig von der Mentalität des Probanden wird so trotz gleichen Sachverhalts in dem einen Fall kein Punkt und im anderen mit einer merklich von Null verschiedenen Wahrscheinlichkeit ein Punkt vergeben.

Sotscheck hat die Zahl der Antwortalternativen mit dem Argument auf sechs reduziert, diese seien leicht zu übersehen und würden so für alle Probanden das gewünschte gleiche Vorwissen sicherstellen. In der Praxis scheinen die Probanden dieser Strategie aber nicht zu folgen. Sie warten, bis sie den Stimulus gehört haben und suchen dann in der Liste. Falls also die Antwortmöglichkeiten auf dem Bildschirm so dargestellt werden, dass die gesuchte Antwort leicht gefunden werden kann, dann kann man die Beschränkung auf sechs Alternativen fallen lassen und alle Alternativen anzeigen. Damit räumt man dem Probanden die Möglichkeit ein, das anzugeben, was er gehört hat. Mit einem Test dieser Art wurde die Qualität der Stimme von Patienten untersucht, die nach Laryngektomie mit Hilfe einer Stimmprothese sprachen (Büttner 1996). Kollmeier (1992) überarbeitete den Test von Sotscheck mit dem Ziel, ihn in der Sprachaudiometrie einsetzen zu können. Fastl et al. (1990) haben den Test

ebenfalls modifiziert, um damit die Akustik von Räumen zu beurteilen. Damit der Einfluss des Nachhalls erfasst werden konnte, boten sie statt eines einzigen Zielworts drei unmittelbar aufeinanderfolgende Wörter an.

In der Fernmeldetechnik fanden Sprachtests, welche als Testitems sogenannte Logatome verwenden, weite Verbreitung. Da die Logatome aus sinnlosen Silben bestehen, lassen sich alle denkbaren Häufigkeitsverteilungen und Lautkombinationen leicht realisieren. In der Tatsache, dass beim Hören der sinnlosen Silben die sonst aktiven komplexen Verarbeitungsprozesse, die auf die zahlreichen gespeicherten Muster zurückgreifen, ins Leere laufen, kann man auch einen Vorteil für audiologische Untersuchungen sehen (Döring, Hamacher 1992, Dölp 1992, Welge-Lüssen 1997).

Einen anderen Ansatz verfolgte Heller (1992). Er führte die frequenzspezifische Sprachaudiometrie ein und benutzte als Sprachmaterial Wortpaare, die sich in nur einem Konsonanten unterscheiden. Die maximale spektrale Energiedichte dieser Konsonanten lag bei unterschiedlichen Frequenzen. Die Ergebnisse liefern dadurch einen Hinweis, welche Frequenzen für die Schwerhörigkeit verantwortlich sind.

Im Allgemeinen kommunizieren Menschen im Alltag nicht durch den Gebrauch einzelner Wörter, sondern mit sinnvollen Sätzen. Nach Niemeyer (1967) ist der Satz die kleinste rhetorische Einheit, mit der man einen vollständigen Gedanken äußern kann, weshalb Sätze die im Alltag am häufigsten benutzte sprachliche Ausdrucksform darstellen. Wenn man also die Kommunikationsfähigkeit eines Menschen im Alltag erfassen möchte, so sollte man dies auf der Grundlage von Sätzen tun, folgerte Niemeyer und arbeitete einen entsprechenden Test, den Marburger Satztest, aus. Er besteht aus 8 Listen mit je 10 kurzen Sätzen von insgesamt 50 Wörtern (Niemeyer, Beckmann 1962, Niemeyer 1967). Die Sätze sind unterschiedlich lang und umfassen vier bis sechs Wörter. Die Listen wurden unter Mitwirkung von Helmut Meier weitgehend ausbalanciert. Die Verständlichkeit des Marburger Sprachtests liegt zwischen der für die Freiburger Zahlen und der für die Einsilber (Niemeyer 1967, Brinkmann 1974). Dieser Test zeigt eine Innenohrschwerhörigkeit deutlicher an als der Freiburger Einsilber-Test und erweist sich daher als hilfreich zur integralen Beurteilung der

auditiven Kommunikationsfähigkeit (Tschopp, Probst 1992). Eine überarbeitete Variante des Marburger Tests stellt der Göttinger Satztest dar (Wesselkamp 1992, Kollmeier, Wesselkamp 1997).

Die erwähnten Sprachtests werden den Patienten oder Probanden häufig nicht in Ruhe dargeboten, sondern unter Störschallbedingungen. Dafür lassen sich mehrere Gründe anführen:

- a) Es fiel auf, dass viele Schwerhörige vor allem über Probleme beim Hören in lauter Umgebung klagen, ohne dass sich diese Schwierigkeit durch den Verlauf der Hörschwelle im Tonaudiogramm erklären ließe (v. Wedel 1985, Pekkarinen et al. 1990, Tschopp, Probst 1992). Daraus folgt, dass eine Untersuchung des Patienten ausschließlich in ruhiger Umgebung nicht genügt, um auf seine Kommunikationsfähigkeit unter Alltagsbedingungen zu schließen.
- b) Die eben angesprochene Erfahrung, dass klinische Tests in schallisolierten Räumen nicht in jedem Fall auf den Alltag verallgemeinert werden können, ließ den Wunsch aufkommen, die Testsituation dem Alltag anzunähern. Und da im Alltag eine Unterhaltung selten in völliger Ruhe stattfindet, sondern in einer mehr oder weniger lauten Umgebung, sollten diese Störgeräusche bei der Untersuchung nachgeahmt werden (Niemeyer 1976).
- c) Die Sprachverständlichkeit ist begrenzt auf das Intervall zwischen 0 % und 100 %. Falls zwei Probanden in Ruhe alle Wörter verstehen, so kann der Test zwischen diesen beiden Probanden nicht differenzieren, obwohl sie über eine durchaus unterschiedliche Fähigkeit, Sprache zu verstehen, verfügen können. Diese sogenannten Decken- und Bodeneffekte kann man vermeiden, wenn man der Sprache Störgeräusche überlagert. Es wird gerade soviel Rauschen hinzugefügt, dass eine vorgegebene Sprachverständlichkeit, meist 50 %, erreicht wird. Im zugehörigen Signal-Rausch-Verhältnis, der sogenannten Sprachverständlichkeitsschwelle, gewinnt man ein alternatives Maß, um das Hörvermögen eines Probanden zu charakterisieren. Dieses ist vorteilhafterweise nicht wie die Sprachverständlichkeit in seinem Wertevorrat beschränkt. (Martin 1997, Hagerman 1984).

- d) Die Sprachverständlichkeitsfunktionen verlaufen im Störschall steiler als in Ruhe (Kießling et al. 1994). Dies bringt einen ökonomischen Vorteil: Die Sprachverständlichkeit kann ohne Genauigkeitseinbuße mit weniger Testlisten geschätzt werden.
- e) Letztlich beeinflusst das Rauschen abhängig von der Art der Schwerhörigkeit das Frequenz- und Pegelunterscheidungsvermögen sowie das Zeitauflösungsvermögen des Gehörs in unterschiedlicher Weise (Hahlbrock 1970, (Schorn, Zwicker 1986).

Die Verständlichkeit von Sätzen unterliegt noch einem weiteren Effekt. Im Kontext eines Satzes werden mehr Wörter verstanden als wenn diese isoliert angeboten werden (Bosman 1989, Züst, Tschopp 1995). Bekanntermaßen kann man sich völlig ungestört unterhalten, auch wenn man nicht jedes Phonem klar und deutlich perzipiert. Selbst unter ungünstigen Bedingungen stören die Lücken kaum, weil auf Grund phonologischer, syntaktischer und semantischer Regeln sowie lexikalischer Einschränkungen nur wenige Items als mögliche Kandidaten übrigbleiben. Die Trefferwahrscheinlichkeit ist bei einem kleinen vorgegebenen Vorrat natürlicherweise größer als bei einem großen. Miller et al. (1951) haben diesen Effekt experimentell nachgewiesen. Kalikow und Stevens (1977) konstruierten Sätze, in denen ein Zielwort leicht oder schwer (high or low predictability) vorherzusagen war, und maßen, wie sich dieser Unterschied auf die Verständlichkeit des Zielwortes auswirkte. Mehrere Autoren haben Modelle entwickelt, um diesen Einfluss des Kontextes quantitativ fassen zu können. Shannon (1951) wendete den von ihm entwickelten informationstheoretischen Ansatz auf das Englische an. Der durch den Satz vermittelte Kontext bildet sich in einer verminderten linguistischen „Entropie“ ab. Boothroyd und Nittrouer (1988) lassen sich vom Multiplikationssatz der Wahrscheinlichkeiten leiten. Als bestimmend für die Satzverständlichkeit fungiert als Parameter eine „effektive“ Wortzahl. Diese ist in einem Satz ohne Kontext gleich der Anzahl der Wörter im Satz. In einem Satz mit Kontext ist sie umso kleiner, je größer der Kontext ist. Bronkhorst et al. (1993, 2002) formulierten ein alternatives, zweistufiges Modell auf der Grundlage expliziter Wahrscheinlichkeiten für das Erkennen der

Elemente eines Satzes auf der ersten Stufe und das Ersetzen nicht erkannter Elemente auf der zweiten Stufe. Das Modell kann an experimentelle Daten angepasst werden und liefert so die gewünschten Kontextinformationen.

Da die Verarbeitung von Kontexteffekten zentralen Prozessen zugeschrieben wird, lagen Versuche nahe, Kontexteffekte für die Untersuchung von zentralen Hörstörungen zu nutzen (Tschopp, Probst 1992, von Werra et al. 1995).

2.3 Diskutierte Nachteile von Sprachtests

Manche Autoren halten es für einen Nachteil der Satztests, dass die gemessene Verständlichkeit nicht unerheblich davon abhängt, wie geschickt der Proband die gehörten Wörter zu einem sinnvollen Ganzen kombinieren kann. Beispielsweise argumentiert Langenbeck (1963) in dieser Weise. Wenn es um den Sprachtest als diagnostisches Instrument geht, dann sieht er in dem Einfluss der Kombinationsfähigkeit des Patienten einen die Varianz vergrößern und deshalb störenden Faktor. Gleiches gilt für eine unterschiedlich ausgeprägte Fähigkeit, auf Wortschatz und grammatikalische Regeln zugreifen zu können.

Insgesamt ist der Prozess des Sprachverstehens wegen seiner Komplexität schwer fassbar. Entsprechend vielfältig sind die theoretischen Ansätze zu seiner Beschreibung: Autonomiehypothese (Fodor 1983), Interaktionshypothese (Marslen-Wilson, Tyler 1980), Logogen-Modell (Morton 1969), das Cohort-Modell I (Marslen-Wilson, Tyler 1980) und Cohort-Modell II (Marslen-Wilson 1987), Trace-Modell (McClelland, Elman 1986) u. a. Bei der Vielzahl der Hypothesen und Modelle überrascht es nicht, dass teilweise unterschiedliche Ziele verfolgt wurden und sich manche Betrachtungsweisen konträr gegenüberstehen. So vertritt Niemeyer (1967) die zu Langenbeck (1963) gegensätzliche Meinung, dass Schwerhörige Sprache in der gleichen Weise verarbeiten wie Normalhörige. Die Kombinationen auf der Satzebene wären eine integrierende Teilfunktion des Sprachverständnisses (Niemeyer, Beckmann 1962). Der Sprachprüfung mit Sätzen könne man keine übernormale Verständlichkeit vorwerfen. Die Testpersonen greifen lediglich auf die normalen, dauernd

wirksamen Prozesse des natürlichen Sprachverständnisses zurück. Durch eine Prüfung mit Einzelwörtern könne diese nicht erfasst werden. Allerdings konzidiert Niemeier, dass auch ein Interesse an der Prüfung zentraler Hörleistungen bestünde, wofür spezielle Tests, beispielsweise der „Binaural Test“ von Matzker (1958) und Feldmann (1965), zur Verfügung ständen.

Falls man es für erwiesen ansieht, dass die Verständlichkeitsmessung mit unterschiedlichen Tests unterschiedliche Ergebnisse und ebenso unterschiedliche Informationen liefert, dann liegt es nahe, mehrere Tests zu einer ‚Test-Batterie‘ zusammenzufassen. Dieser Ansatz wurde in den Vereinigten Staaten zuerst aufgegriffen. Als Beispiel für eine deutschsprachige Spielart sei auf den MAC-Test (minimum auditory capabilities) von Dillier und Spillmann (1992) verwiesen. Gedacht ist der Test für die präoperative Untersuchung und das postoperative Management von CI-Kandidaten.

2.4 Rauschen

Es wurde schon gesagt, dass Menschen gewöhnlich nicht in vollkommener Stille miteinander kommunizieren, sondern meist in einer geräuschvollen Umgebung. Aus untersuchungstechnischer Sicht gibt es leider viel zu viele unterschiedliche Geräusche, von leisen Umweltgeräuschen angefangen, leiser Musik oder geflüsterter Stimme bis zu lautem Schreien, lauten Motorgeräuschen oder noch lauterem Donner. Gerade Schwerhörige bekommen in etwas lauterer Umgebung Probleme, einer Unterhaltung folgen zu können. Daher hat die Forderung, vor allem wenn der Nutzen von Hörhilfen beurteilt werden soll, die sprachaudiometrischen Untersuchungen im Störlärm durchzuführen, inzwischen Eingang ins Lehrbuch gefunden (Kießling 1996).

Trotz einer Vielzahl von störenden Lärmquellen werden für die audiologischen Routineuntersuchungen nur wenige, typische Geräusche eingesetzt:

A) Weißes Rauschen: Weißes Rauschen zeichnet sich durch ein frequenzunabhängiges Leistungsdichtespektrum und eine gausförmige Amplitudenverteilung aus. Subjektiv wird es als höhenlastiges Zischen wahrgenommen, da das menschliche Gehör bei etwa 2 kHz seine größte Empfindlichkeit besitzt.

B) Rosa Rauschen: Das Leistungsdichtespektrum des Rosa Rauschens nimmt pro Oktave um 3 dB ab. Da die Hörschwelle zu den tiefen Frequenzen hin ab der Stelle größter Empfindlichkeit ansteigt und damit verbunden die Lautheit von Signalen gleichen Pegels abnimmt, klingt Rosa Rauschen in der Tonhöhe ausgeglichen.

C) Rauschen nach CCITT Rec. G.227: Dieses Rauschen bildet die spektralen Eigenschaften gesprochener Sprache nach. Mit Rauschen nach CCITT lässt sich daher Sprache effizient verdecken. Da bei vielen Umweltgeräuschen ein Frequenzbereich um einige hundert Hertz betont ist, ähnlich wie dies auch bei Sprache der Fall ist, wird Rauschen nach CCITT auch gern benutzt, um eine alltägliche Geräuschkulisse im Labor zu simulieren. Die Zeitstruktur gesprochener Sprache wird damit jedoch nicht nachgebildet. Vorteilhafterweise lässt sich dieses Rauschen einfach durch geeignete Filter aus weißem Rauschen erzeugen.

D) Rauschen nach Fastl: Dieses Rauschen hat die gleiche spektrale Energieverteilung wie das Rauschen nach CCITT Rec. 227. Es approximiert aber gleichzeitig die Lautstärkeschwankungen gesprochener Sprache durch eine Amplitudenmodulation. Laut Fastl (1987) liegt dieses Störgeräusch hinsichtlich der Beeinträchtigung der Wortverständlichkeit zwischen den beiden Extremen, nämlich einem einzigen konkurrierenden Sprecher und einem Stimmengewirr.

E) Sprache: Häufig wird die Sprachverständlichkeit beeinträchtigt, wenn viele Personen durcheinander reden. Sprache als Störgeräusch lässt sich unterschiedlich konfigurieren. Zwischen den oben angesprochenen Extremen gibt es beliebige Übergänge von zwei, drei usw. konkurrierenden Sprechern. Falls sehr viele Personen gleichzeitig sprechen, so spricht man von einem Stimmengewirr. Die größte zeitliche Fluktuation erhält man mit einem Sprecher, die kleinste mit dem Stimmengewirr. Letzteres gleicht im Zeitverhalten einem gefilterten weißen Rauschen.

F) Als Störsignal sind auch Musik, fremde Sprache und vieles mehr verwendet worden. Die Möglichkeiten scheinen unbegrenzt zu sein. Unter gewissen Aspekten mögen solche Störgeräusche sinnvoll erscheinen, für die sprachaudiometrische Routine haben sie keine Bedeutung erlangt.

Bezüglich der spektralen Energieverteilung des Rauschsignals ist zu beachten, dass die verdeckende Wirkung über den Frequenzbereich des Rauschens hinausgeht (Zwicker, Fastl 1999, v. Wedel 1985). Ein schmalbandiges Signal maskiert zwar erwartungsgemäß maximal jenen Frequenzbereich, der für das Signal konstituierend ist, aber es verdeckt auch in niedrigeren und effizienter in höheren Frequenzbereichen. Mit zunehmender Intensität des Rauschens nimmt die maskierende Wirkung im höherfrequenten Bereich überproportional zu. Wegen der Gestalt der Hörfläche kann ein leise empfundenes, tieffrequentes Rauschen hochfrequente Bereiche merkbar maskieren.

2.5 Anforderungen an den Aufbau eines Satztests

Die Sprachtests sollen bei einer mit einem CI versorgten Person aufzeigen, wie gut diese mit dem CI versteht, welche Unterschiede zu anderen bestehen und wie sich das Sprachverständnis mit der Zeit entwickelt. Damit dies mit geringem Aufwand möglich ist, muss der Test möglichst reproduzierbare Ergebnisse liefern, d. h. Ergebnisse mit kleiner Varianz. Alle Varianz erzeugenden Effekte sollten daher möglichst ausgeschaltet werden. Im Allgemeinen unterscheiden sich Menschen in Bezug auf Bildung und Lebenserfahrung. Auch diese Unterschiede sollten die Testergebnisse möglichst wenig beeinflussen. Deshalb sollten Wortschatz und Satzbau dem bekannten Gebrauch der Alltagssprache folgen, wie Niemeyer (1967) forderte. Regional verwendete Wörter sollten vermieden werden, gleichfalls Namen, wozu auch Ortsnamen zählen, ebenso wissenschaftliche Begriffe, Tabu-Wörter usw. (Dillier, Spillmann 1992). Die Sätze sollten auch ein ausgewogenes Maß an Redundanz enthalten. Außerdem sollten sie nicht zu schwierig oder zu verwirrend sein. Die Sätze sollten auch nicht zu lang sein, damit sie der Proband sicher wiederholen kann (Plomp, Mimpfen 1979).

Die unterschiedlichen Listen sollten untereinander äquivalent und phonologisch der Sprache angepasst sein. Sie sollten alle möglichen Manifestationen einer Sprache widerspiegeln. Konsequenterweise sollte der Test auch affirmative und interrogative Sätze enthalten. Außerdem sollte ein Test so umfangreich sein,

dass bei wiederholten Tests die gleichen Sätze nicht mehrfach angeboten werden müssen, wodurch die Gefahr besteht, dass sich die Testperson an bereits gehörte Sätze erinnert. Darüber hinaus sollten die Listen möglichst kurz und handlich sein, damit die Untersuchung nicht zu viel Zeit in Anspruch nimmt und damit auch ökonomischen Zwängen gerecht wird. Ferner sollten die Probanden während der Untersuchung nicht ermüden. Eine nachlassende Konzentration und Aufmerksamkeit würde die Ergebnisse ebenfalls verfälschen. Die genannten Anforderungen sind nicht alle gleichzeitig zu erfüllen, so dass die Umsetzung zu mehr oder weniger großen Kompromissen zwingt.

Eine Entscheidung ist auch hinsichtlich der Art zu treffen, wie der Sprachtest ausgewertet werden soll:

- a) Man kann die richtig wiedergegebenen Sätze zählen. Ein Satz gilt als richtig verstanden, wenn jedes Wort im Satz richtig verstanden wurde. Diese Art der Auswertung gilt als weniger effektiv als die Auszählung der richtigen wiedergegebenen Wörter (Versfeld 2000).
- b) Man kann nur die richtig wiedergegebenen Wörter zählen. Dies ist die am häufigsten angewandte Methode.
- c) Man kann auf die richtig wiedergegebenen Phoneme abstellen.
- d) Und schließlich kann man nur darauf achten, ob ein in den Satz eingebettetes Schlüsselwort richtig verstanden wurde. Man unterscheidet zwischen Sätzen, in denen das Schlüsselwort einfach vorherzusehen ist (High Predictability) und solchen, in denen dies schwierig ist (Low Predictability) (Kalikow, Stevens 1977, Niklès, Tschopp 1996).

2.6 HSM-Satztest

Wie bereits angeführt, wurde der HSM-Satztest mit dem Ziel entwickelt, die Sprachverständlichkeit bei CI-Trägern zu messen (Hochmair-Desoyer et al. 1997). Der Name leitet sich von den Autoren Hochmair, Schulz und Moser ab. Der Test enthält eine ausreichende Anzahl von Sätzen, nämlich 30 Listen mit je 20 Sätzen. Jede Liste enthält 14 Aussagesätze und sechs Fragesätze. Die Anzahl der Wörter im Satz bewegt sich zwischen drei und acht. Alle Listen sind

strukturgleich aufgebaut und jede Liste besteht aus 106 Wörtern. Die Sätze wurden, wie dies für einen Satztest gefordert wird, der Alltagssprache entnommen. Dialektwörter und lokale Besonderheiten wurden vermieden.

Um gleiche Verständlichkeit der Listen zu erreichen, wurde eine Vorversion mit Normalhörenden bei einem festen Signal-Rausch-Abstand von -10 dB untersucht. Aus dem umfangreichen Ausgangsmaterial konnten die Listen unter Einhaltung der angegebenen Nebenbedingungen so zusammengestellt werden, dass sie gleich verständlich waren. In einer nachfolgenden Untersuchung (Schmidt 1997) wurde die Abhängigkeit der Diskriminationsfunktion vom Signal-Rausch-Pegel ermittelt. Die Sprachverständlichkeitsschwelle, das ist derjenige Signal-Rausch-Abstand, bei welchem genau 50 % richtig verstanden werden, lag bei -7,2 dB. Die Steilheit der Diskriminationsfunktion an der Sprachverständlichkeitsschwelle betrug 10 %/dB.

Bei CI-Patienten lag die Sprachverständlichkeitsschwelle bei 14 dB Signal-Rausch-Abstand, die Verständlichkeit nahm an der Schwelle mit 6 %/dB zu. Die Diskriminationsfunktion verläuft bei CI-Patienten also flacher als bei Normalhörenden.

Der Test liegt auf CD vor und ist käuflich zu erwerben. Die Sprache belegt die eine Stereospur, der Störschall, sprachsimulierendes Rauschen nach CCITT, Rec. 227, die andere.

2.7 In Griechenland gebräuchliche Sprachtests

Im griechischen Sprachraum werden für die Sprachverständlichkeitsmessungen zweistellige Zahlen und zweisilbige Wörter verwendet. Einsilbige Wörter kommen im Griechischen nur selten vor, daher gibt es kein dem Freiburger Einsilber entsprechendes Pendant. Bisläng gab es auch keinen griechischen Satztest. Einen Zahlen-Test haben Manolidis und Kastellis unabhängig voneinander entwickelt (Manolidis 1964, Kastellis 2009). Tests mit zweisilbigen Wörtern haben Kogias, Manolidis und Kastellis vorgestellt (Kogias 1961, 1974, Manolidis 1964, Kastellis, Battmer 1978). Die Wortlisten sind phonetisch ausbalanciert und untereinander äquivalent.

Da eine umfassende Auszählung der Phoneme des Griechischen fehlte, haben die oben genannten Autoren die Vokale und Konsonanten in einigen neugriechischen Texten ausgezählt und die so gewonnenen statistischen Ergebnisse für den Aufbau ihrer Tests benutzt. Im Druck gleich erscheinende, aber unterschiedlich ausgesprochene Konsonantenkomplexe wurden als verschieden behandelt. Manolidis hat in seinen Tests auch die Phonemumgebungen berücksichtigt. Im Altgriechischen verwendete man zur Unterscheidung der Vokale prosodische Merkmale, diese Unterschiede gibt es im Neugriechischen nicht mehr. Neuere Tests mit phonemisch ausbalancierten Listen zu je 50 zweisilbigen Wörtern wurden von Iliadou et al. (2006) und Trimmis et al. (2006) vorgestellt. Letztere bedienten sich der modernen Compact Disc (Trimmis et al. 2007).

Neuerdings wurde eine groß angelegte Untersuchung durchgeführt, in der aus 48 000 Texten 33 000 000 Wörter extrahiert wurden (Hatzigeorgiu et al. 2000, Gavrilidou et al. 1999, 2002, Mikros et al. 2005). Iliadou et al. stützten die Auswahl der Wörter bereits auf die Arbeiten des Teams von Hatzigeorgiu, Gavrilidou et al. über die Häufigkeitsverteilung der Wörter im Griechischen ab.

2.8 Ziel der Arbeit

Da in Griechenland ein dem HSM-Satztest entsprechender Test bislang fehlte, aber an mehreren Kliniken eine wachsende Zahl von Patienten mit Cochlea-Implantaten versorgt wird, schien es angebracht, auch für das Griechische einen solchen Test zu entwickeln. Daraus leitet sich die Zielsetzung dieser Arbeit ab, nämlich die aufgezeigte Lücke zu schließen und einen Sprachtest nach dem Vorbild des HSM-Satztests zu entwickeln. Die Sätze sollten wie beim HSM-Test der Alltagssprache entstammen. Außerdem sollten auch die anderen, weiter oben ausgeführten Anforderungen an einen Sprachtest möglichst vollständig erfüllt werden.

3 Vorbereitende Arbeiten

3.1 Materialsammlung

Zu Beginn der Arbeit wurde eine Sammlung von griechischen Alltagssätzen angelegt. In diesen Sätzen sollte sich möglichst der Gebrauch des Griechischen im normalen Alltag widerspiegeln. Im Griechischen benutzt man wie im Deutschen unterschiedlich lange Sätze. Die angelegte Sammlung enthielt daher Sätze unterschiedlicher Länge und unterschiedlicher Satzart. Wie beim HSM-Test waren es Aussage- und Fragesätze. Die Anzahl der Wörter pro Satz variiert zwischen drei und elf. Insgesamt umfasste diese primäre Satzsammlung 914 Sätze.

In einem nächsten Schritt wurden die Sätze in Athen auf CD gesprochen. Die Auflösung betrug 16 Bit und die Abtastfrequenz 44 kHz. Das Aufnahme-Equipment bestand aus einer Revox MB 16 Konsole, einem Rode Broadcast Mikrophon und einem Sony DTSZE 700 Dat Recorder. Die Aufzeichnung fand in einer schallisolierten Kabine statt. Herr Moser, HNO Würzburg, schlug einen geschulten, männlichen Rundfunksprecher im Alter zwischen 35 und 50 Jahren vor. Der ausgewählte Sprecher, Hr. Siakawellas, erfüllte mit einem Alter von 38 Jahren die Vorgaben bezüglich des Alters und des Geschlechts. Er wurde gebeten, sich bei der Aufsprache um eine gleichbleibende Lautstärke zu bemühen. Die Lautstärkeschwankungen sollten sich innerhalb von ± 2 dB bewegen. Da diese Toleranzgrenze schwer einzuhalten ist, wurde jeder Satz hintereinander zweimal aufgesprochen. Bei der späteren Auswahl der Sätze konnte man auf diese Weise auf den besseren Satz zurückgreifen.

In einem nächsten Schritt sichtete und überarbeitete Herrn M. Schmidt in Innsbruck die Aufsprache mit der Software Cool Edit Pro. Er brachte alle Sätze auf den gleichen maximalen Schalldruckpegel. Von den beiden gleichen Sätzen übernahm er den mit der besser ausgeglichenen Aufsprache. Wegen zu großer Lautstärkedifferenzen einzelner Wörter in beiden Sätzen wurden 203 Sätze als nicht geeignet eingestuft und ausgesondert. Für die nachfolgend durchgeführten Sprachverständlichkeitsmessungen verblieben 711 Sätze.

3.2 Generierung des Rauschsignals

Die vorgesehenen Untersuchungen wurden mit einem speziellen Störgeräusch durchgeführt. Zunächst wurden die Sätze bei der Überarbeitung in Innsbruck zu einem Stimmengewirr übereinander kopiert. Da ein gesprochener Satz eine deutliche Zeitstruktur der Lautstärke zeigt, erzeugt Sprache als Rauschsignal eine zusätzliche Varianz im Testergebnis. Um den Vorteil einer angepassten spektralen Energieverteilung des Stimmengewirrs mit der eines zeitlich wenig fluktuierenden Signals verbinden zu können, wurde weißem Rauschen das Spektrum des erhaltenen Stimmengewirrs aufgeprägt. Das Spektrum des so generierten ‚sprachähnlichen‘ Rauschens gleicht damit vollständig dem der im Test benutzten Sätze und spiegelt auch die stimmlichen Charakteristiken des Sprechers wider. Es folgt aber einer anderen Zeitstruktur.

Das Rauschsignal wurde zu jedem Satz in passender Länge auf den zweiten Stereokanal der CD aufgezeichnet. Als Referenz für die Lautstärke der Signale dient beim Rauschen dessen Effektivwert, bei der Sprache der für alle Sätze gleiche maximale Pegel innerhalb eines Satzes. Die angegebenen Werte für das jeweils benutzte Signal-Rausch-Verhältnis beziehen sich auf diese beiden Pegel. Das Rauschen wird mit einer Rampe von 200 ms angefahren, und nach weiteren 200 ms beginnt die Wiedergabe des Satzes. Am Ende des Satzes wird das Rauschsignal mit einer gleichfalls 200 ms langen Rampe ausgeblendet. Die relativ lange Vorlaufphase für das Rauschen soll sicherzustellen, dass sich die AGCs der CIs bereits im eingeschwungenen Zustand befinden, wenn die Sprache einsetzt.

4 Vorstudie

Die Vorstudie sollte bei sinnvollem Aufwand vorab eine Information über die Verständlichkeit der aufgezeichneten Sätze liefern, um die Listen in der Hauptstudie mit einem Signal-Rausch-Verhältnis anbieten zu können, das zu einer mittleren Verständlichkeit führt.

4.1 Material und Methoden

Die Untersuchung stützt sich auf 6 Probanden, 4 Männer und 2 Frauen. Alle Probanden waren gebürtige, junge Griechen, die in Griechenland aufgewachsen waren und in Würzburg studierten. Alle sprachen Griechisch als Muttersprache und waren 21 bis 23 Jahre alt. Das Durchschnittsalter betrug 22,5 Jahre. Die Probanden waren leicht für die Untersuchung zu motivieren.

Vor Durchführung des Tests wurden die Ohren inspiziert und eine Stimmgabeluntersuchung durchgeführt. Anschließend wurde ein Tonaudiogramm aufgenommen. Kein Proband hatte in einer der getesteten Frequenzen von 250 bis 8 000 Hz einen Hörverlust größer als 15 dB. Alle Probanden wiesen in der Anamnese keine Erkrankung oder Operation der Ohren auf. Eine erblich bedingte Schwerhörigkeit war nicht bekannt.

Aus den in Innsbruck nach der Überarbeitung verbliebenen 711 Sätzen wurden 20 Listen zu je 35 Sätzen gebildet. Das ist insgesamt ein Vorrat von 700 Sätzen. Da man davon ausgehen musste, dass auch strukturgleiche Sätze nicht gleich verständlich sein würden, wurde für diese Voruntersuchung keine strikte Strukturgleichheit der Listen angestrebt. Ziel der Untersuchung war es, die Verständlichkeit eines jeden einzelnen Satzes zu messen und mit diesen Informationen die endgültigen Listen zusammenzustellen. Die für diese Untersuchung benutzten Listen hatten daher vorläufigen Charakter und vereinfachten Planung und Durchführung. Die Anzahl der Wörter pro Liste variierte von 208 bis 215, so dass beim Abhören trotzdem eine ungefähr gleich große Belastung erreicht wurde. Die Häufigkeitsverteilung ist in Abb. 1 wiedergegeben. Am häufigsten vertreten waren Listen mit 209 Wörtern.

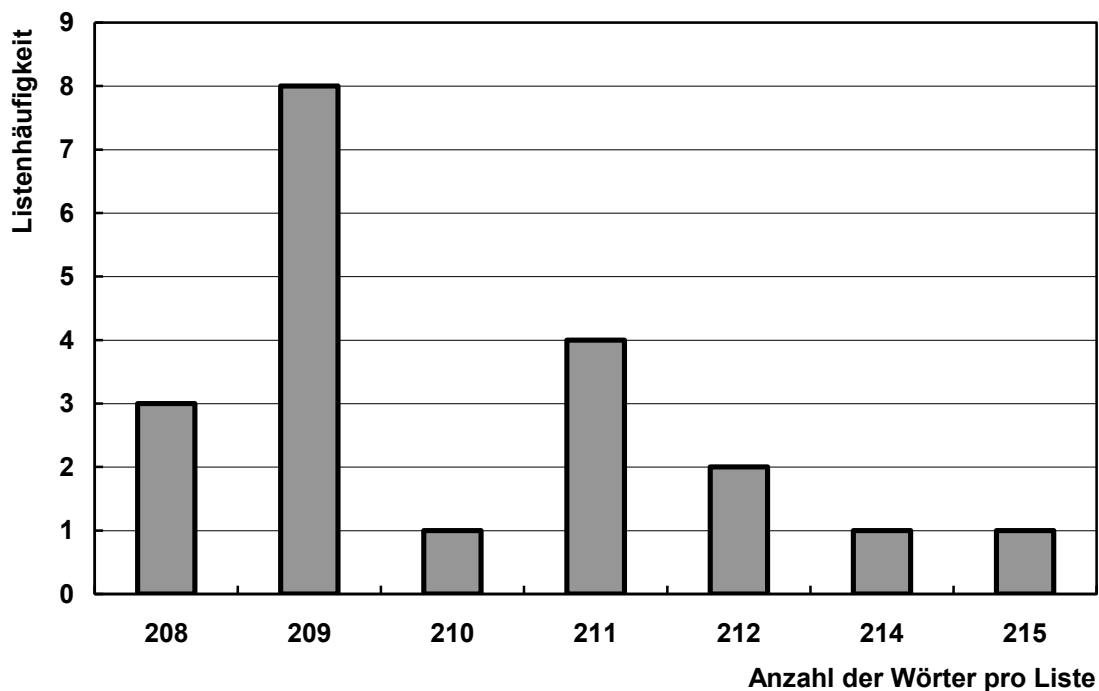


Abb. 1: Häufigkeitsverteilung der Listen aufgeschlüsselt nach der Anzahl der enthaltenen Wörter

4.2 Versuchsdurchführung

Alle Messungen fanden in der HNO-Klinik der Universität Würzburg statt. Die Probanden hörten die Sprachproben in einer geräumigen, schallisolierten Kabine ab. Sprache und Rauschen wurden über den gleichen Lautsprecher von vorne angeboten. Der Abstand zwischen Lautsprecher und Probanden betrug einen Meter, wie dies Lehnhardt et al. (1996) vorschlagen. Die Lautstärke der Sätze und des Rauschens wurde mit Hilfe des Audiometers CAD-03 der Firma Westra eingestellt. Sprach- und Rauschsignale wurden zusammengemischt und über eine einzige Audiometriebox Lab-1001 des gleichen Herstellers wiedergegeben.

Vor Beginn der Untersuchung wurde mit dem Schallpegelmesser Typ 2231 der Firma Brüel & Kjaer sichergestellt, dass die Lautstärke in Kopfhöhe am Ort des Probanden mit der Anzeige des Audiometers übereinstimmte.

Jeder Proband hörte 8 Listen bei vier unterschiedlichen Signal-Rausch-Abständen ab. Die Sprache wurde bei allen Testdurchgängen bei einem konstanten Pegel von 65 dB dargeboten, während das Rauschen im Pegel verändert wurde. Das ist die Lautstärke, in der sich normalerweise zwei Personen unterhalten (Lehnhardt 1996, Niemeyer 1976). Die vier Signal-Rausch-Abstände betragen 0 dB, -2 dB -4 dB und -6 dB. Bei einem negativen Signal-Rausch-Abstand ist das Rauschen lauter als die Sprache. Alternativ könnte man auch den Sprachpegel variieren. Dadurch erhielte man aber schwieriger zu interpretierende Ergebnisse, weil sich die Sprachverständlichkeit sowohl mit Lautstärke der Sprache also auch mit dem Signal-Rausch-Abstand verändert.

Begonnen wurde stets mit einem Signal-Rausch-Abstand von 0 dB. Die nachfolgenden drei Pegel wurden randomisiert angeboten. Gesteuert wurde die gesamte Testdurchführung von einem in Matlab erstellten Programm. Die Sätze der aktuell vorgespielten Liste wurden auf dem zugehörigen Monitor in einem Fenster angezeigt. Der jeweils angebotene Satz war in der Darstellung hervorgehoben.

Die Probanden sollten die Sätze, so weit sie diese verstanden hatten, nachsprechen. Die richtig nachgesprochenen Wörter wurden auf dem Bildschirm mit der Maus markiert und automatisch gezählt. Am Ende eines jeden Durchgangs zeigte das Audiometer den Prozentsatz der richtig verstandenen Wörter an und speicherte die erhobenen Daten in einer Datenbank.

Die Testlisten enthielten auch zusammengesetzte Wörter. Diese wurden bei der Auszählung folgendermaßen behandelt. Aus zwei Hauptwörtern zusammengesetzte Begriffe wurden wie zwei getrennte Wörter gewertet. Falls das zusammengesetzte Wort eine Präposition enthielt und diese mehr als fünf Silben umfasste, wurden die Präposition und das restliche Wort als einzelne Wörter gezählt. Insgesamt wurde jede Liste zweimal getestet, wobei keine Liste zweimal mit dem gleichen SNR abgehört wurde.

Eine Sitzung dauerte ungefähr 2 Stunden, weswegen die Probanden zwischendurch eine kleine Pause einlegen durften. Die Vorgehensweise war in Hauptstudie und Vorstudie identisch.

4.3 Ergebnisse

Wie die untenstehende Tabelle zeigt, steigt die mittlere Sprachverständlichkeit mit wachsendem Signal-Rausch-Abstand von 21 % auf 82 % an. Dieses Intervall schließt, wie gewünscht, den 50-%-Wert ein, welcher mit der Sprachverständlichkeitsschwelle assoziiert ist.

| | 0 dB | -2 dB | -4 dB | -6 dB |
|------------|------|-------|-------|-------|
| Proband A | 72 | 53 | 36 | 14 |
| Proband B | 79 | 66 | 36 | 36 |
| Proband C | 86 | 79 | 51 | 32 |
| Proband D | 88 | 68 | 45 | 19 |
| Proband E | 91 | 55 | 48 | 15 |
| Proband F | 75 | 70 | 25 | 11 |
| Mittelwert | 82,2 | 65,4 | 40,5 | 21,4 |

Tabelle 1: Mittelwerte der Sprachverständlichkeit in der Vorstudie aufgeschlüsselt nach Probanden und Gesamtmittelwert.

5 Hauptstudie

5.1 Material, Methoden, Versuchsdurchführung

An der Hauptstudie haben 20 Probanden im Alter zwischen 21 und 28 Jahren teilgenommen, 12 Frauen und 8 Männer. Das Durchschnittsalter betrug 24,5 Jahre. Die Probanden der Hauptstudie erfüllten die gleichen Kriterien wie die Probanden der Vorstudie.

Die Hauptstudie wurde in der gleichen Weise wie die Vorstudie durchgeführt: Die gleiche schallisolierte Kabine, das gleiche Instrumentarium, der gleiche Abstand zwischen Proband und Lautsprecher von einem Meter, Sprache und Rauschen kamen wiederum aus einem einzigen Lautsprecher, nur wurden jetzt alle 20 Listen zu je 35 Sätzen verwendet. Auch hier speicherte das Audiometer die als richtig oder falsch markierten Wörter in eine Datenbank, welche für die weitere Auswertung nach Excel übertragen werden konnte.

Jeder Proband hörte zunächst übungsweise zwei Listen mit 0 dB und -4 dB Signal-Rausch-Abstand ab. Bei 20 Probanden und 20 Listen wurde jede Liste zweimal als Übungsliste verwendet, alternierend einmal als erste und einmal als zweite Liste. Die Ergebnisse aus diesen beiden Übungssequenzen flossen nicht in die Auswertung ein. Anschließend hörte jeder Proband 10 Listen bei einem Signal-Rausch-Abstand von -4 dB ab. Um die Belastung für die Probanden in Grenzen zu halten, wurden in der Hauptstudie die Listen nur bei einem einzigen Signal-Rausch-Abstand angeboten. Der Wert von -4 dB wurde aus den Ergebnissen der Vorversuche abgeleitet.

Wenn die Probanden längere Zeit konzentriert der Untersuchung folgen sollen, so ist zu erwarten, dass sie einerseits mit der Zeit ermüden und die Treffer abnehmen, und dass sie andererseits mit der Versuchsdurchführung zunehmend vertrauter werden, sich besser in den Test einhören und die Zahl der richtig verstandenen Wörter steigt. Um den dadurch bewirkten systematischen Verzerrungen in den Ergebnissen entgegenzuwirken, wurden die benutzten Listen von Messung zu Messung und von Proband zu Proband systematisch variiert. Proband 1 sollte mit Liste 1 beginnen, gefolgt von Liste 2 und schließlich mit Liste 10 enden. Proband 2 sollte als Erstes die Liste 2 abhören und

dann die zehn nachfolgenden Listen bis Liste 11. Jeder folgende Proband sollte seine Sitzung mit der jeweils nächsten Liste beginnen. Da nur 20 Listen zur Verfügung standen, lag es nahe, die Listen zyklisch aneinanderzureihen. Falls im Laufe der Präsentationen die Liste 20 erreicht würde, so sollte die Reihe wieder mit Liste 1 beginnend fortgeschrieben werden. Die daraus resultierende vollständige Versuchsmatrix ist in Tabelle 2 wiedergegeben. Bei der Umsetzung verschob sich auf Grund eines Übertragungsfehlers bei drei Probanden die Anfangsposition.

Da nicht alle Probanden alle Listen abhörten, liegt ein hierarchischer Versuchsplan vor, und weil jeder Proband mehrere Listen abgehörte, eine Messwieder-

| Proband | Liste 1 | Liste 2 | Liste 3 | Liste 4 | Liste 5 | Liste 6 | Liste 7 | Liste 8 | Liste 9 | Liste 10 | Liste 11 | Liste 12 | Liste 13 | Liste 14 | Liste 15 | Liste 16 | Liste 17 | Liste 18 | Liste 19 | Liste 20 |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | | | | | | | | |
| 2 | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | | | | | | | |
| 3 | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | | | | | | |
| 4 | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | | | | | |
| 5 | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | | | | |
| 6 | | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | | | |
| 7 | | | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | | |
| 8 | | | | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | |
| 9 | | | | | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | |
| 10 | | | | | | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| 11 | | | | | | | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 12 | 10 | | | | | | | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 13 | 9 | 10 | | | | | | | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 14 | 8 | 9 | 10 | | | | | | | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 15 | 7 | 8 | 9 | 10 | | | | | | | | | | | 1 | 2 | 3 | 4 | 5 | 6 |
| 16 | 6 | 7 | 8 | 9 | 10 | | | | | | | | | | | 1 | 2 | 3 | 4 | 5 |
| 17 | 5 | 6 | 7 | 8 | 9 | 10 | | | | | | | | | | | 1 | 2 | 3 | 4 |
| 18 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | | | | | | | | | 1 | 2 | 3 |
| 19 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | | | | | | | | | 1 | 2 |
| 20 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | | | | | | | | | | 1 |

Tabelle 2: Geplante Reihenfolge, in der die Probanden die Listen abhören sollten.

holung auf dem Faktor Proband. Der verwendete hierarchische Versuchplan unterstellt, dass den Interaktionen nur eine geringe Bedeutung zukommt und der wesentliche Zusammenhang durch die Haupteffekte beschrieben wird. Das zugrunde gelegte Modell behauptet daher, dass die Verständlichkeit V einer Liste einem für alle Listen gleichen Niveau M folgt und durch additiv wirkende Effekte modifiziert wird. Es werden drei Effekte in das Modell einbezogen. Möglicherweise sind die Listen nicht gleich gut verständlich. Für jede Liste L_i $\{i = 1..20\}$ wäre der Mittelwert M der Verständlichkeit dann um einen Wert dL_i nach oben oder unten zu korrigieren. Die gleichen Überlegungen für den Probanden P_j $\{j = 1..20\}$ angestellt, führt auf die Korrektur dP_j , und ein Lerneffekt würde abhängig von der Position k $\{k = 1..10\}$ in der zeitlichen Abfolge der Listen die Korrektur dA_k bedingen.

Verständlichkeitsmessungen sind wie alle Messungen nicht exakt reproduzierbar, sondern zeigen ein gewisses Maß an Variabilität. Sie gelten deshalb als fehlerbehaftet. Das eben skizzierte Modell ist daher noch um einen Fehlerterm zu erweitern. Da im vorliegenden Fall eventuell vorhandene Interaktionen in diesen Term mit einfließen, wird hierfür die Bezeichnung Residuum $R_{i,j,k}$ gewählt. Mit den eingeführten Bezeichnungen ergibt sich folgendes Modell:

$$V(L_i, P_j, A_k) = M + dL_i + dP_j + dA_k + R_{i,j,k}$$

Die Probanden können als eine Stichprobe aus der Grundgesamtheit aller griechisch sprechenden Menschen mit ähnlichen Merkmalen angesehen werden. Sie werden daher wie ein zufälliger Faktor behandelt. Die Nullhypothese bezüglich der Probanden lautet daher, die Varianz der $dP(k)$ verschwindet. Die Listen sind Ergebnis eines Konstruktionsprozesses. Sie fungieren daher als feste Effekte. Die zeitliche Reihenfolge in der Darbietung der Listen ist ebenfalls wie ein fester Effekt zu behandeln.

Beim Import der Daten aus der Audiometrie-Datenbank nach Excel entstand ein inhaltlich und strukturell gleicher Datensatz. Jede Zeile enthält in mehreren Feldern die charakterisierenden Merkmalsausprägungen und den zugeordneten Wert für die gemessene Verständlichkeit. Die Werte der oben benannten Variablen bilden sich daher als Spaltenvektoren ab. Die zeitaufwendige und wegen der großen Anzahl auch fehleranfällige Eingabe der Formeln für die

üblicherweise benutzten Mittelwerte wurde umgangen, stattdessen wurden mit Hilfe der Funktion ‚Solver‘ die Komponenten so bestimmt, dass die Summe der Residuenquadrate minimiert und die gesuchten Variablen orthogonalisiert wurden. Jeder der oben erwähnten Effekte wurde in einer zusätzlichen Spalte zurückgegeben.

Für jede Effekt-Spalte und die Residuen-Spalte wurden die Quadrate der Einzelwerte aufsummiert und durch die Freiheitsgrade geteilt. Mit der F-Verteilung wurden die Effekte an der Residuenvarianz auf Signifikanz geprüft. Das Signifikanzniveau wurde einheitlich auf 5 % festgelegt. Da die benutzte Residuenvarianz Interaktionen einschließt, fällt die Entscheidung konservativ aus.

Aus den vorliegenden Daten wurde für jeden Satz die Standardabweichung der Verständlichkeit berechnet. Diese ist für die 700 Sätze in Abb. 7 dargestellt. Die Punktwolke scheint sich trotz der ersichtlich großen Streuung auf einen ellipsenförmigen Bereich oberhalb der Abszisse zu erstrecken. Einen einfachen, im vorliegenden Fall allerdings nur approximativen Zugang zur vorliegenden Problematik erlaubt die Theorie der linearen Funktionen unabhängiger Zufallsvariablen. Man wird die Unabhängigkeit als ausreichend gegeben annehmen dürfen, wenn aus der Verständlichkeit der Sätze auf die der Listen geschlossen wird, da in Sprachtests die Sätze beziehungslos aufeinanderfolgen. Anders sieht es bei den Wörtern in einem Satz aus. Diese sind sehr eng aufeinander bezogen, und das Ereignis, ein Wort zu verstehen, ist keineswegs unabhängig davon, wie die anderen Wörter im Satz verstanden werden. Aber es ist durchaus aufschlussreich zu sehen, wie sich Verletzungen der Voraussetzungen auf die erhobenen Daten auswirken.

Für das Weitere wird das Verstehen eines Satzes oder eines Wortes als ein stochastischer Prozess betrachtet, der mit einer gewissen Wahrscheinlichkeit ein Ereignis zeitigt. Verschiedene Ereignisse sollen zunächst als voneinander unabhängig angenommen werden.

Als Maß für die Sprachverständlichkeit diene die Anzahl der richtig verstandenen Wörter. Diese Art, den Sprachtest auszuwerten, lässt sich in die Wahrscheinlichkeitstheorie einbetten, indem man, anstatt zu zählen, für jedes richtig

verstandene Wort die Zahl 1 und für jedes nicht verstandene Wort die Zahl 0 aufaddiert. Formal geschieht dies, indem man für jedes Wort i eine Zufallsvariable K_i , definiert, welche die Werte 1 oder 0 annimmt, je nachdem ob das Wort verstanden wurde oder nicht. Die Zufallsvariable K als Summe der K_i „zählt“ dann die richtig verstandenen Wörter im Satz ab, wobei n für die Satzlänge, die Anzahl der zu beurteilenden Wörter steht:

$$K = K_1 + K_2 + \dots + K_n$$

Die Variable K_i nimmt die Werte 1 oder 0 mit den Wahrscheinlichkeiten p_i und $q_i = 1 - p_i$ an. Für jeden Satz lagen mehrere gemessene Werte der Verständlichkeit vor, woraus als repräsentativer Wert der arithmetische Mittelwert der richtig wiedergegebenen Wörter berechnet wurde. Dem entspricht im stochastischen Modell der Erwartungswert von K , nämlich $E(K)$. Wegen der Linearität des Erwartungswertes addieren sich die Erwartungswerte $E(K_i)$ für die Einzelereignisse zu $E(K)$ auf:

$$\begin{aligned} E(K) &= E(K_1) + E(K_2) + \dots + E(K_n) \\ &= p_1 + p_2 + \dots + p_n \end{aligned}$$

Dieses Ergebnis ist eine verwertbare Beziehung, welche den Mittelwert auf die Wahrscheinlichkeiten zurückführt. In gleicher Weise addieren sich auch die Erwartungswerte für die quadratischen Abweichungen, die Varianzen. Die Varianz für das Wort i ergibt sich aus der Definitionsgleichung des Erwartungswertes zu $p_i \cdot (1 - p_i)$.

$$\begin{aligned} V(K) &= V(K_1) + V(K_2) + \dots + V(K_n) \\ &= p_1(1 - p_1) + p_2(1 - p_2) + \dots + p_n(1 - p_n) \end{aligned}$$

Die Satzverständlichkeiten sind per definitionem auf das Intervall zwischen 0 % und 100 % beschränkt. Die Werte 0 % und 100 % können nur beobachtet werden, wenn entweder keines oder alle Wörter im Satz verstanden werden. In beiden Fällen verschwindet also die Varianz. Die oben angegebenen Modellgleichungen führen zu dem gleichen Ergebnis. $E(K) = 0$ oder $E(K) = n$ bedingt $p_i = 0$ oder $p_i = 1$ für alle i aus $\{1..n\}$. Daraus folgt nach der angegebenen Beziehung $V(K) = 0$.

Der Term $p_i(1 - p_i)$ beschreibt eine Parabel, welche für $p_i = 0$ und 1 die Abszisse schneidet und für $p_i = \frac{1}{2}$ das Maximum bei $\frac{1}{4}$ durchläuft. Falls alle p_i den

gleichen Wert $p = \frac{1}{2}$ annehmen, dann liegt ein Bernoulli-Experiment vor. Natürlich beträgt in diesem Fall auch die mittlere Satzverständlichkeit 50 %, und die Varianz $V(K)$ nimmt ihren maximalen Wert von $n/4$ an. Nicht richtig ist allerdings die Umkehrung, dass nämlich für eine Satzverständlichkeit von 50 % die Varianz $V(K)$ ihren maximalen Wert von $n/4$ annehmen muss. $E(K)$ stellt sich ja als Summe von n unabhängigen Werten dar, die sich in mancherlei Weise aufaddieren können. Im Extremfall kann für die eine Hälfte der p_i gelten: $p_i = 0$, und für die andere Hälfte: $p_i = 1$. Die Werte addieren sich dann zu $E(K) = n/2$ auf und die Satzverständlichkeit beträgt 50 %, aber die Varianz verschwindet, $V(K)$ ist gleich Null. Im Allgemeinen werden die p_i irgendwelche Zwischenwerte annehmen. Für die Werte der gemessenen Varianz der einzelnen Sätze wird man daher erwarten dürfen, dass sie in regelloser Weise einen Bereich zwischen der Abszisse und einer parabelähnlichen oberen Grenze mit einem Maximum bei $n/4$ überdecken. Bleibt die Frage, was geschieht, wenn nicht länger angenommen werden kann, dass die p_i unabhängig sind. Diese Besonderheit wird später mit Rückgriff auf die Ergebnisse diskutiert werden.

5.2 Ergebnisse

Alle im Modell spezifizierten Haupteffekte erwiesen sich als hoch signifikant. Die Wahrscheinlichkeiten für einen Fehler erster Art ergaben sich für den Beitrag der Probanden zu $p = 3 \cdot 10^{-35}$, für die Listen zu $p = 3 \cdot 10^{-61}$ und für Reihenfolge in der Darbietung zu $p = 6 \cdot 10^{-4}$. Die in den Messwerten vorhandene Varianz wird zu 87 % durch die Variablen des Modells erklärt. Abb. 2 veranschaulicht den Zusammenhang zwischen den Vorhersagen des linearen Modells und den gemessenen Werten. Idealerweise würden alle Werte auf der eingezeichneten

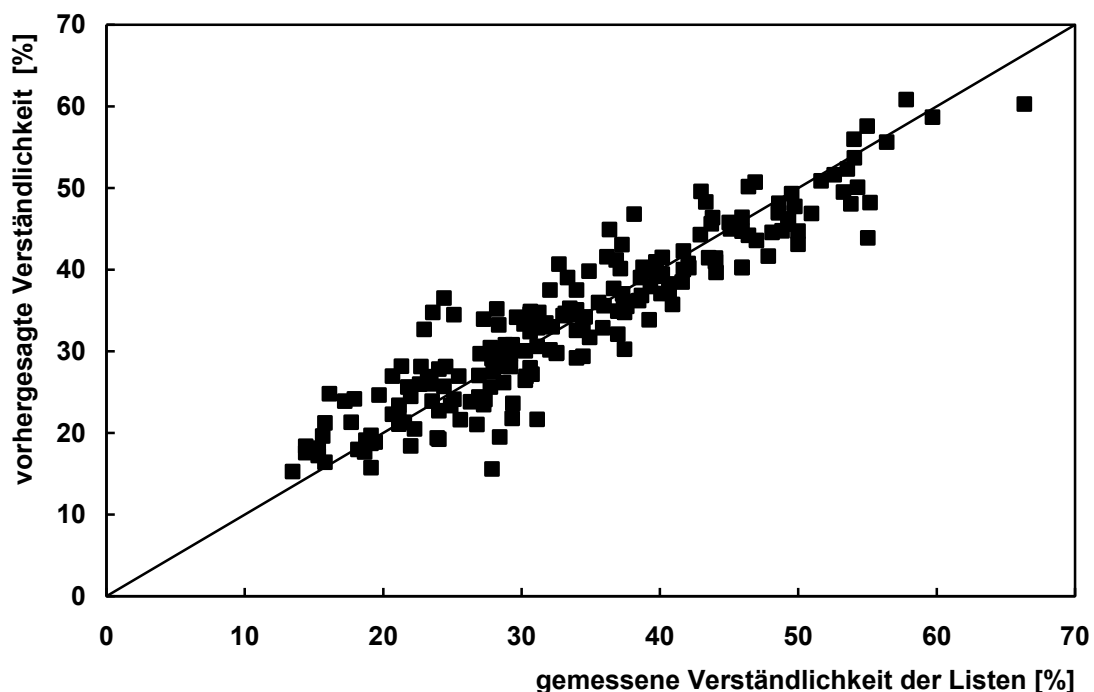


Abb. 2: Vergleich der gemessenen Verständlichkeit mit der durch das Modell vorhergesagten Verständlichkeit

Diagonalen liegen. Der quadratisch gemittelte Vorhersagefehler beträgt 4,1 %. Die Nullhypothese, die Residuen sind normal verteilt, lässt sich bei einem $p = 0.62$ im Shapiro-Wilk-Test auf Normalverteilung nicht verwerfen. Der Mittelwert der Verständlichkeit liegt bei 34,2 %.

Wie Abb. 3 zeigt, verstehen manche Probanden besser als andere. Charakterisiert man jeden Probanden durch den Mittelwert der prozentual verstandenen Wörter und sortiert danach, so verteilen sich die Probanden relativ gleichmäßig

über den Bereich von 25 % bis 45 % Verständlichkeit. Männer und Frauen verstehen gleich gut, sie unterscheiden sich nicht signifikant in den Modellparametern.

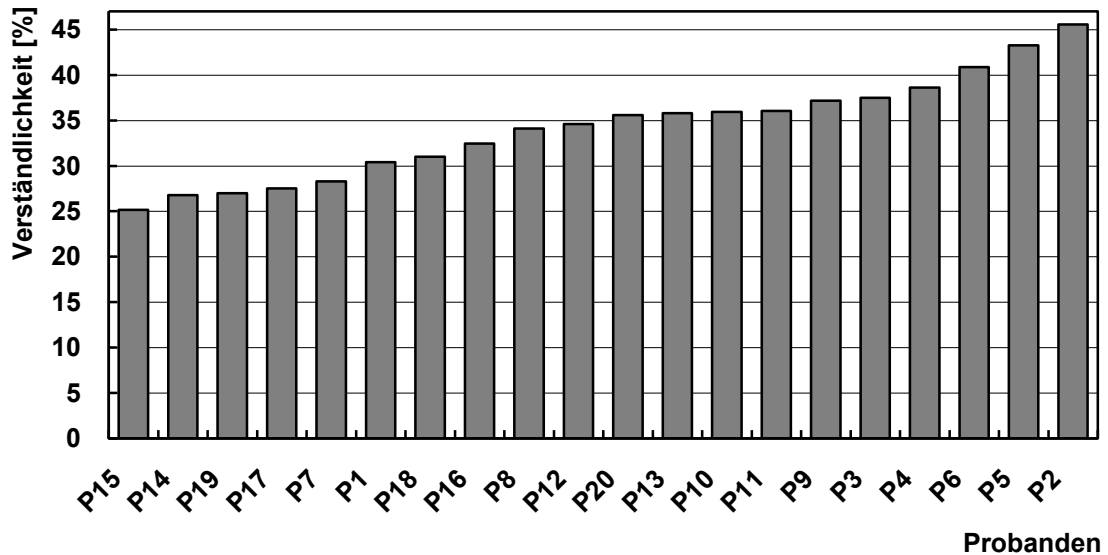


Abb. 3: Um Listen- und Lerneffekte korrigierte Verständlichkeit der Testlisten für die einzelnen Probanden

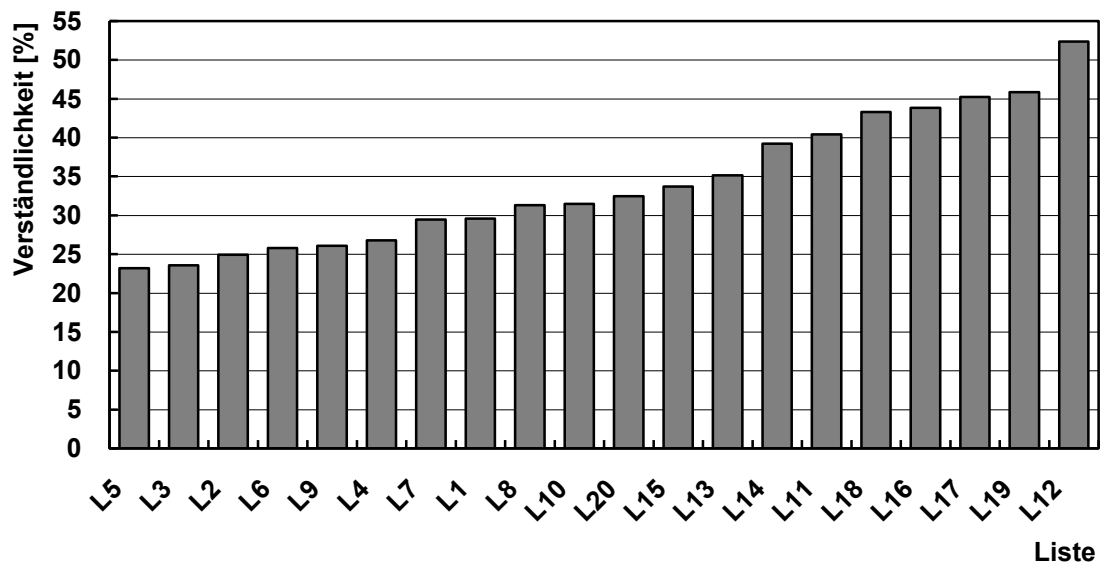


Abb. 4: Um Probanden- und Lerneffekte korrigierte Verständlichkeit der einzelnen Testlisten

Ähnliches gilt für die unterschiedliche Verständlichkeit der Listen. Hier reichen die Verständlichkeitswerte von 23 % bis 52 %. Abb. 4 zeigt die Details.

Über die 10 in Folge abgehörten Listen hinweg errechnet sich ein Lerneffekt von etwa 4 %. Die Einzelwerte und die Regressionsgerade sind in Abb. 5 dargestellt. Natürlich ist nicht anzunehmen, dass der Lerneffekt, wie dies die Regressionsgerade nahe legen könnte, linear mit der Zeit fortschreitet. Plausibel ist eher ein zunehmend flacher werdender Verlauf, der sich asymptotisch einem oberen Grenzwert annähert.

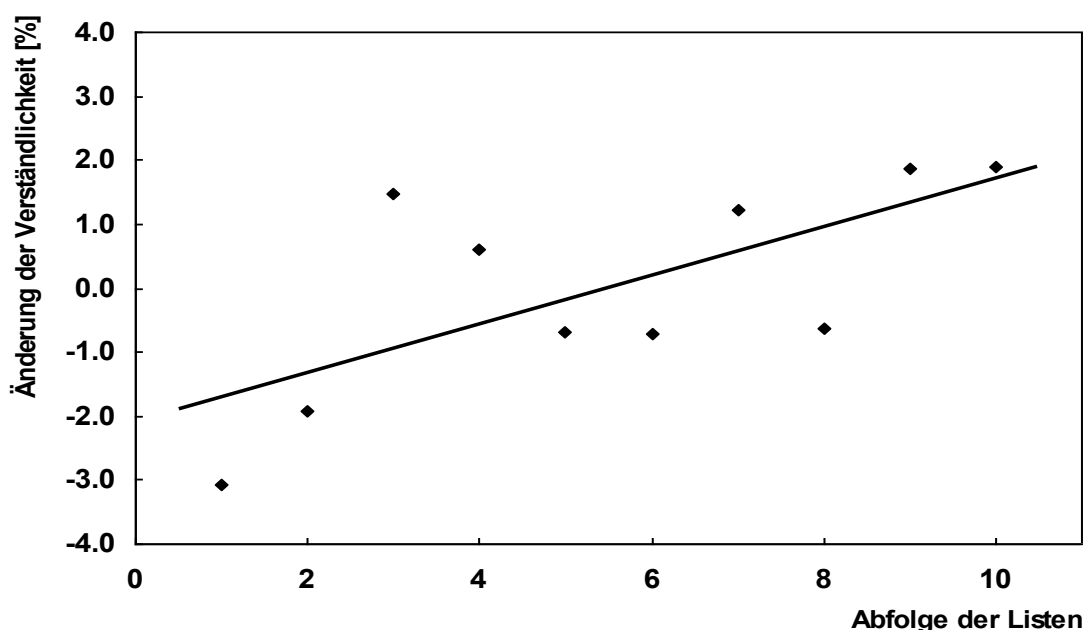


Abb. 5: Abhängigkeit der Verständlichkeit von der Darbietungsreihenfolge

Obwohl Sprache im Allgemeinen nicht so erlebt wird, offenbaren Sprachtests eine immanente Inhomogenität. Das Phänomen unterschiedlicher Verständlichkeiten zeigt sich schon bei den kleinsten Bausteinen einer Sprache, den Phonemen (Büttner 1996). Sie pflanzt sich über Silben, Wörter und Sätze bis in die Listen fort. Abb. 6 zeigt die unterschiedliche Verständlichkeit aller Sätze unter Verwendung des ‚Gauß Papiers‘. Die einzelnen Sätze sind extrem unterschiedlich verständlich. In manchen Sätzen wird kein einziges Wort verstanden, in anderen dagegen alle Wörter. Wegen des deutlich sichtbaren Boden- und

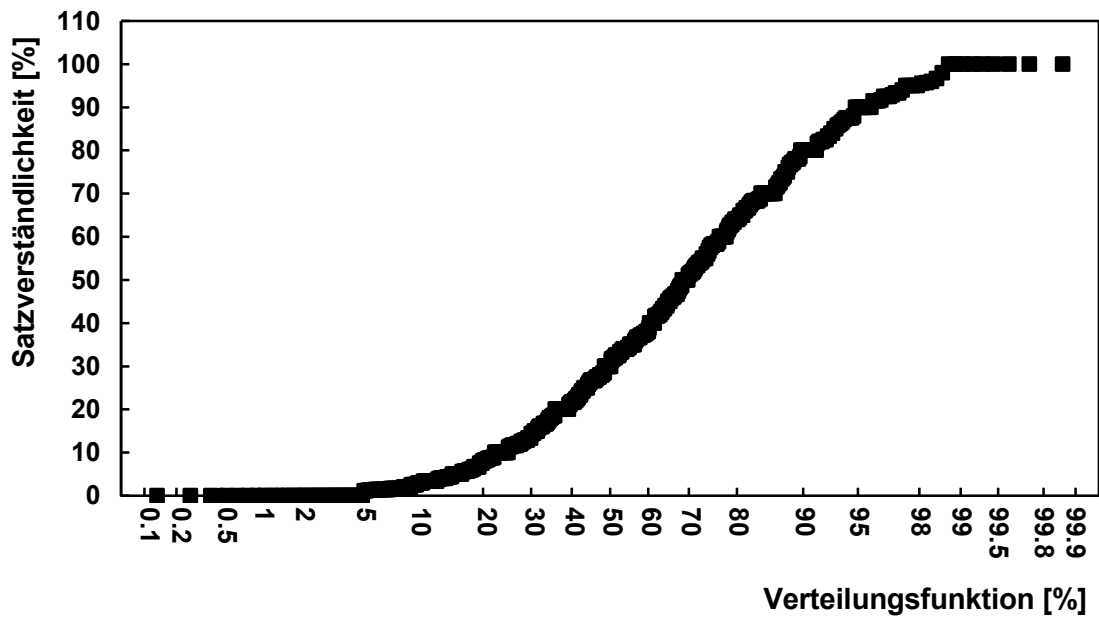


Abb. 6: Kumulierte Häufigkeitsverteilung der gemessenen Verständlichkeit der einzelnen Sätze

Deckeneffekts ist die Verständlichkeit der einzelnen Sätze nicht normal verteilt. Bei einer Normalverteilung lägen die Werte näherungsweise auf einer Geraden. Die Konsequenzen, die sich aus dieser Verteilung für die Konstruktion des Tests ergaben, werden im nächsten Kapitel untersucht.

Eine wichtige Eigenschaft eines Messinstruments, hier des Satztestes, spiegelt sich in der damit erzielbaren Genauigkeit wider. Im Folgenden wird als Genauigkeitsmaß die aus den Messwerten berechnete Standardabweichung benutzt. Vorteilhafterweise hat sie die gleiche Dimension wie die zugrundeliegenden Messwerte. Für die Darstellung im vorhergehenden Abschnitt schien wegen der Eigenschaft der Additivität dagegen die Varianz besser geeignet. Der Zusammenhang zwischen beiden Größen ist einfach: Die Standardabweichung ist als die Wurzel aus der Varianz definiert. Für die 700 untersuchten Sätze zeigt Abb. 7 die gefundenen Messwerte. Auf der Abszisse ist der Mittelwert der Verständlichkeit aus den jeweils 10 Messungen für jeden Satz und auf der Ordinate die daraus berechnete Standardabweichung aufgetragen. Auf den ersten Blick mag die große Streuung der Messwerte überraschen. Zunächst einmal gilt jedoch für die Stichprobe wie für das stochastische Modell, dass die Streuung für die Extremwerte der Verständlichkeit, nämlich für 0 %

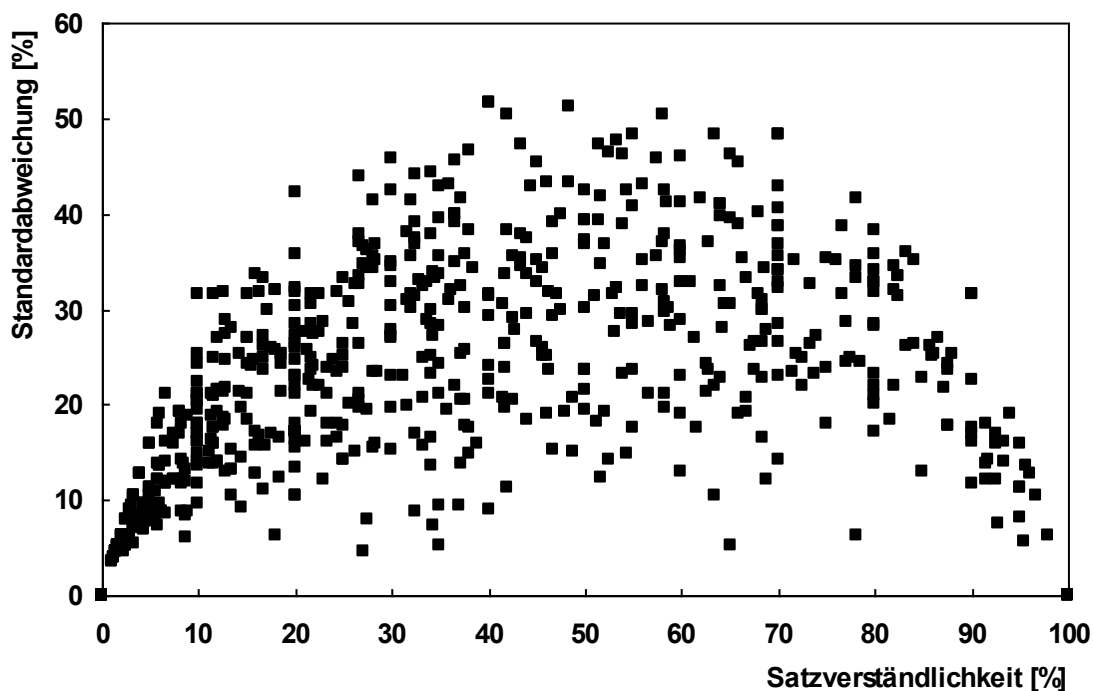


Abb. 7: Standardabweichung aller 700 gemessener Sätze in Abhängigkeit von der Verständlichkeit der einzelnen Sätze

und 100 %, verschwindet. Eine mittlere Verständlichkeit von 0 % kann nur gemessen werden, wenn in der betrachteten Stichprobe kein einziger Satz verstanden wurde. Folglich errechnet man für die Standardabweichung den Wert Null. Gleiches gilt für eine mittlere Verständlichkeit von 100 %, nur dass hier jeder Satz mit 100% verstanden werden muss. Die größten Werte für die gemessene Standardabweichung findet man im Bereich mittlerer Verständlichkeiten. Gleichzeitig streuen in diesem Bereich die dargestellten Werte am meisten. Die Grenzen hierfür lassen sich leicht angeben. Den größtmöglichen Wert nimmt die Standardabweichung dann an, wenn die Abweichung vom Mittelwert für alle Sätze maximal wird. Für eine mittlere Verständlichkeit der Sätze von 50 % bedeutet dies, dass die Hälfte der Sätze gar nicht, die andere Hälfte zu 100 % verstanden wird. Die Summe der Abweichungsquadrate beträgt dann bei jeweils 10 Sätzen in der Stichprobe $10 \cdot (50\%)^2$, und bei 9 Freiheitsgraden berechnet sich daraus $s_{\max} = 52,7\%$. Tatsächlich können mehrere Werte in diesem Größenbereich beobachtet werden, z. B. bei dem

Satz „I polli dulia troi ton afenti“. In diesem Satz haben fünf Probanden kein Wort richtig verstanden, vier alle 6 Wörter und ein Proband 5 Wörter.

Der kleinste Wert, den die Standardabweichung annehmen kann, ist Null. Dieser Fall tritt ein, wenn alle zehn Sätze der Stichprobe zufälligerweise von jedem Probanden mit der gleichen Verständlichkeit abgehört werden. Dabei können die verstandenen Wörter an unterschiedlichen Positionen im Satz auftreten. Insgesamt gibt es also $2^6 = 64$ möglichen Anordnungen, falls man die Anzahl der verstandenen Wörter und deren Position systematisch variiert. Mit einer „1“ für ein verstandenes und einer „0“ für ein unverstandenes Wort könnte die Antwort eines Probanden beispielsweise so (010010) kodiert werden. Es gibt genau jeweils eine Anordnung (000000) und (111111), sechs Anordnungen mit einem richtig verstandenen Wort, welche von (000001) bis (100000) reichen, usw. Die meisten Anordnungen gibt es mit drei verstandenen und nicht verstandenen Wörtern, nämlich $6!/3!/3! = 20$. Die Messergebnisse zeigen ein davon abweichendes Bild. In den Daten finden sich 6 Sätze mit einer Verständlichkeit zwischen 48,7% und 51,3%. Diese sechs Sätze repräsentieren eine kleine Stichprobe von Sätzen mit einer mittleren Verständlichkeit von etwa 50%. Da jeder Satz zehnmal abgehört wurde, enthält die Stichprobe 60 gemessene Werte für die Verständlichkeiten. Von diesen 60 Werten entfallen 14 auf Fälle, in denen kein Wort oder alle Wörter im Satz verstanden wurden. Nach den obigen Überlegungen würde man bei 60 Werten mit etwa zwei Sätzen rechnen. Die „mittlere“ Gruppe mit jeweils drei richtig oder falsch verstandenen Wörtern im Satz bringt es dagegen nur acht Messergebnisse anstelle der erwarteten 20. Diese so andere Häufigkeitsverteilung ist ein deutlicher Hinweis darauf, dass bei Sprachtest die einzelnen Wörter nicht unabhängig voneinander und nicht mit der gleichen Wahrscheinlichkeit verstanden werden. Weitergehendes zu den hier aufscheinenden Besonderheiten in den Daten und zu deren Modellierung durch eine Binomialverteilung findet sich im Abschnitt Diskussion.

Die untersuchten Sätze waren dem alltäglichen Gebrauch entlehnte sinnvolle Sätze. Bei deren Abhören entsteht auf unbekannte Weise entweder gleich oder verzögert vor dem geistigen Auge die Assoziation einer Silbe, eines Worts. Abhängig von Signal-Rauschabstand variiert die Wahrnehmung. Im Extermfall

wird nur ein nicht sinngabendes Geräusch gehört, vielleicht manchmal mit eingestreuten einzelnen Lauten. Mit abnehmendem Rauschen werden immer mehr Wörter erkannt, wovon manche aber innerhalb des Satzes keinen Sinn ergeben. Der menschliche Geist ist bestrebt, eine möglichst sinnvolle Interpretation des Gehörten aufzufinden, weshalb er die einzelnen Lautmuster nicht unabhängig voneinander in ein wahrgenommenes Wort transformiert, sondern mehr oder weniger große Teile eines Satzes während dieses Entscheidungsprozesses ineinander verwebt. Es mag daher nicht überraschen, dass die

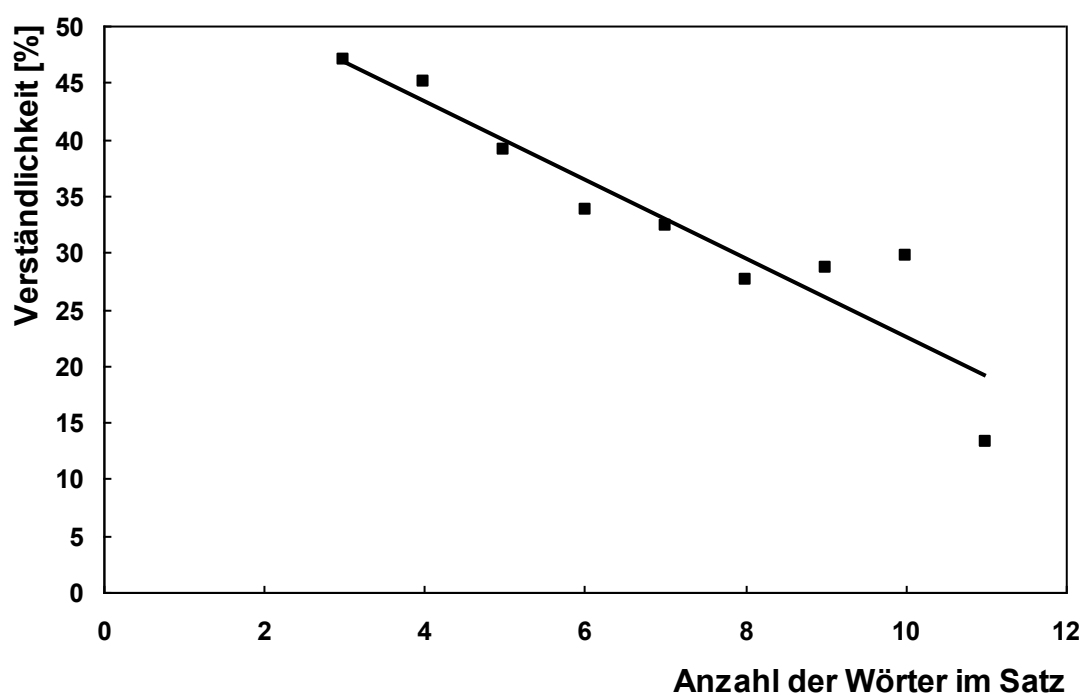


Abb. 8: Abhängigkeit der Satzverständlichkeit von der Anzahl der Wörter, aus denen sich der Satz zusammensetzt

Verständlichkeit eines Satzes, wie Abb. 8 erkennen lässt, auch von dessen Länge, der Anzahl der Wörter im Satz, abhängig ist. Die kurzen Sätze werden signifikant ($p = 2,4 \cdot 10^{-4}$) besser verstanden als die langen. Die Verständlichkeit eines Wortes in einem Dreiwort-Satz ist mehr als doppelt so groß wie in einem Satz mit elf Wörtern.

6 Konstruktion der endgültigen Listen

In beiden Untersuchungen verstanden die Probanden die einzelnen Listen unterschiedlich gut. Wünschenswert sind aber Listen, die gleich gut verständlich sind. Um dies zu erreichen, wurden aus dem vorhandenen Vorrat an Sätzen neue Satzlisten zusammengestellt.

Wie Abb. 6 ausweist, wurden einige Sätze extrem gut und andere extrem schlecht verstanden. Es gibt Sätze, bei denen alle Probanden entweder kein Wort oder alle Wörter im Satz verstanden haben. Sätze dieser Art enthalten nur eine sehr geringe Information über die Lage der Sprachverständlichkeitsschwelle. Sie lassen nur den Schluss zu, dass die Sprachverständlichkeitsschwelle bei höheren oder niedrigeren Werten liegt, aber sie lassen keine Aussage darüber zu, wie weit die Sprachverständlichkeitsschwelle vom aktuellen benutzten Signal-Rausch-Verhältnis entfernt liegt. Offensichtlich ist es daher von Vorteil, solche Sätze bei der Zusammenstellung der endgültigen Listen nicht zu berücksichtigen.

Da, wie schon öfters erwähnt, die Sprachverständlichkeit innerhalb der einzelnen Individuen variiert und man Boden- und Deckeneffekte vermeiden möchte, wird man versuchen, diejenigen Sätze für den Test auszuwählen, die eine mittlere Verständlichkeit aufweisen. Da es wegen der breiten Verteilung der Verständlichkeit nur wenige Sätze mit genau dieser Eigenschaft gibt, wird man, um nicht zu viele Sätze verwerfen zu müssen, einen Kompromiss zwischen diesen beiden Extremen anstreben. In der vorliegenden Arbeit wurden Sätze, deren Verständlichkeit zwischen 10 % und 90 % lag, für auswählbar gehalten. Sätze mit schlechterer oder besserer Verständlichkeit wurden von der Auswahl ausgeschlossen. Auf Grund dieses Kriteriums wurden 135 Sätze verworfen. Außerdem wurden bei einer nochmaligen Durchsicht der Sätze 25 nicht wohlklingende Sätze aus der Liste gestrichen, so dass für die Auswahl in die neu zu erstellenden Listen 540 Sätze verblieben.

Die Antwort auf die Frage, wie viele Listen sich aus diesem Vorrat bilden lassen, hängt von einigen noch nicht diskutierten Eigenschaften dieses Reservoirs ab. Zum einen sollten die Listen nach dem Vorbild des HSM-Tests zusammengestellt werden, d. h., jeder Satz sollte nur einmal in genau einer

Liste vorkommen. Außerdem sollten alle Listen strukturgleich sein, d. h., alle Listen sollten gleich viele Sätze der gleichen Art und Länge enthalten. Die Anzahl der Listen kann daher die Anzahl der Sätze, die in der kleinsten Gruppe vorhanden sind, nicht übersteigen. Aus diesem Grunde wurden Gruppen, die nur wenige Sätze enthielten, verworfen. Dies betraf die Dreiwort-Sätze und die Sätze mit 10 und 11 Wörtern. Die nächstgrößere Gruppe bildeten die Neunwort-Sätze. Sie umfasst 26 Sätze. Um nicht jeden Spielraum für die Auswahl aus dieser Gruppe zu verlieren, sollte die Anzahl der zu erstellenden Listen etwas kleiner gewählt werden. Letztlich wurde eine Anzahl von 20 Listen für optimal angesehen.

Wie beim HSM-Satztest in seiner ursprünglichen Form sollte jede Liste zehn Sätze enthalten. Die Erweiterung auf 20 Sätze lässt sich, falls gewünscht, leicht verwirklichen, indem man zwei Listen zu 10 Sätzen hintereinander abhört.

Ingesamt setzen sich die Listen aus zwei Fragesätzen und acht Aussagesätzen zusammen: Die beiden Fragesätze waren fünf bzw. sechs Wörter lang. Von den Aussagesätzen war je einer neun, acht und vier Wörter lang. Vier Sätze zählten je sieben bzw. fünf und drei Sätze sechs Wörter.

Neben dieser strukturellen Gleichartigkeit sollten die Listen vor allem gleich verständlich sein. Als Wert für diese allen Listen gemeinsame Verständlichkeit wurden 40 % gewählt. Um alle genannten Nebenbedingungen erfüllen zu können, bedarf es eines überlegten, zielgerichteten Vorgehens. Beispielsweise ist Folgendes leicht vorhersehbar: Man hat neun Sätze ausgewählt und muss nun mit dem zehnten, dem letzten noch auswählbaren Satz, die Zielverständlichkeit erreichen. Falls nun die letzte noch zur Verfügung stehende Gruppe nur wenige Sätze und überdies lange Sätze enthält, dann ist es sehr unwahrscheinlich, dass man in dieser Gruppe noch einen Satz mit der passenden Anzahl richtig verstandener Worte finden kann. Daher wurde der Neunwort-Satz, welcher im Test der längste Satz ist und dessen Gruppe mit nur 26 Sätzen zu den kleinen zählt, als erster Satz ausgewählt.

Die Auswahl der weiteren Sätze erfolgte nach folgender Strategie: Auf jeder Stufe der Auswahl wurde versucht, die Zielverständlichkeit möglichst gut zu treffen. Außerdem wurde eine vorgegebene Toleranzgrenze mit der Auswahl

eines jeden Satzes verringert. Der letzte Satz sollte dann zielgenau ausgewählt werden. Würde man bei diesem Verfahren die Listen immer in der gleichen Reihenfolge bedienen, so würde die erste Liste Sätze nahe der Zielverständlichkeit enthalten und die letzte Liste würde die größten Schwankungen aufweisen. Um dies zu vermeiden, wurde vor jeder Auswahl eines Satzes erst eine Liste ausgelost und dann für diese Liste der passende Satz gesucht.

Wie gerade dargelegt, war es recht einfach, ein Kriterium zu finden, aus welcher Gruppe der erste Satz einer Liste ausgewählt werden sollte. Aber wie ist bei den übrigen Sätzen zu verfahren? Die Problematik zeigt sich am deutlichsten bei der Auswahl des letzten Satzes. Für die Auswahl dieses Satzes sollte man diejenige Gruppe zu Verfügung haben, deren Werte für die Satzverständlichkeit die feinste Körnigkeit aufweisen. Ein Maß hierfür ergibt sich aus der Breite der Verteilung dividiert durch die Anzahl der Sätze in der Gruppe. Den kleinsten Wert hierfür erzielten die Sechswort-Sätze. Das Kriterium lässt sich aber auf alle Gruppen anwenden und impliziert eine Rangfolge für die Auswahl. Diese ließ sich in wenigen Probeläufen nachjustieren und führte zu folgender Reihenfolge: 9, 5?, 6?, 8, 7, 7, 6, 5, 4, 6. (Die Zahlen bezeichnen die Anzahl der Wörter im Satz, das Fragezeichen kennzeichnet die Fragesätze.)

Mit dem dargestellten Auswahlverfahren konnte die Zielverständlichkeit bei 17 Listen exakt erreicht werden, bei drei Listen weicht sie um wenige Promillepunkte ab. Die Abweichungen sind so gering, dass sie für die praktische Anwendung des Tests ohne jede Bedeutung sind.

Als Beispiel für die erzeugten 20 Listen mag die in Tabelle 3 wiedergegebene Liste 1 dienen. Eine vollständige Zusammenstellung aller Listen befindet sich im Anhang.

Liste 1:

- 1 (6 Wörter) To spiti mas echi dio patomata.
- 2 (7 Wörter) Fewgume gia tin ekdromi awrio to proi.
- 3 (5 Wörter) Pion tha psifisete stis ekloges?
- 4 (9 Wörter) Ta kalokeria mas aresun i diakopes konta sti thalassa.
- 5 (4 Wörter) Grafi panta me penna.
- 6 (8 Wörter) Meta apo polli skepsi sinenese sto schedio mas.
- 7 (6 Wörter) Posi zachari thelete ston kafe sas?
- 8 (5 Wörter) To riaki schimatizi pollus katarraktes.
- 9 (7 Wörter) Woithise tin giagia na perasi ton dromo.
- 10 (6 Wörter) I gnosi xenon glosson ine chrisimi.

Κατάλογος 1:

- 1 Το σπίτι μας έχει δύο πατώματα.
- 2 Φεύγουμε για την εκδρομή αύριο το πρωί.
- 3 Ποιόν θα ψηφίσετε στις εκλογές;
- 4 Τα καλοκαίρια μας αρέσουν οι διακοπές κοντά στη θάλασσα.
- 5 Γράφει πάντα με πένα.
- 6 Μετά από πολλή σκέψη συναίνεσε στο σχέδιό μας.
- 7 Πόση ζάχαρη θέλετε στον καφέ σας;
- 8 Το ρυάκι σχηματίζει πολλούς καταρράκτες.
- 9 Βοήθησε την γιαγιά να περάσει τον δρόμο.
- 10 Η γνώση ξένων γλωσσών είναι χρήσιμη.

Tabelle 3: Beispiel einer Liste aus dem Satztest für das Griechische

7 Diskussion

Schwerhörige entwickeln spezifische Techniken, um im Alltag trotz ihres Handicaps zurechtzukommen. Beispielsweise verringern sie den Abstand zu ihren Gesprächspartnern und bringen den Kopf in eine günstige Lauschposition. Zuhause stellen sie Radio und Fernsehgerät laut. Andere versuchen, von den Lippen abzulesen. Solche Verhaltensweisen geben einen Hinweis auf eine vorliegende Hörminderung, eine quantitative Aussage über die Größe des Hörschadens lässt sich daraus aber nicht ableiten. Diesem Ziel dienen audiometrischen Untersuchungsmethoden. Obwohl sich die Tonaudiometrie als Standardverfahren in der Klinik etablieren konnte, bedarf sie bei vielen Fragestellungen der Ergänzung durch die Sprachaudiometrie (Hahlbrock 1970, Niemeyer 1967, Bertolli 1996). Obwohl ein Zusammenhang zwischen Ton- und Sprachaudiogramm zwingend erscheint, ist dieser so vielschichtig, dass bisherige Versuche, diesen herzustellen, zu einer eher nur losen Korrelation führten (Neuss 1982, Dölp 1992, Welge-Lüssen et al. 1997, Smoorenburg 1982, 1992, Brand, Kollmeier 2002).

Weil Sprache ein so komplexes Phänomen darstellt, überrascht es nicht, dass viele Faktoren deren Verständlichkeit beeinflussen. Daraus resultierte die Entwicklung einer Vielzahl von Sprachverständlichkeitstests. Jede Spielart hat spezifische Schwachstellen, die Anlass zu Kritik geben (Dillier, Spillmann 1992, Alich 1985). Der Freiburger Sprachtest macht da keine Ausnahme (Albrecht 1967, Alich 1985, Bangert 1980, Sukowski et al. 2009). Es scheint, dass es einen universell einsetzbaren Sprachtest nicht geben kann. Jede Fragestellung bedarf idealerweise ihres eigenen, angepassten Tests.

Der hier vorgestellte Satztest ist einer speziellen Klasse von audiologischen Tests zuzuordnen, die darauf abzielen, Einschränkungen in der zwischenmenschlichen Kommunikation in einer alltagsnahen Situation zu erfassen. Um eine möglichst wirklichkeitsnahe Testsituation zu erreichen, wurden sinnvolle Sätze, die der Umgangssprache entlehnt sind, als Testmaterial benutzt. Die Verständlichkeit von Sätzen ist hoch, da die Konstruktion eines Satzes festen grammatikalischen Regeln gehorcht und die benutzten Wörter in einem Sinnzusammenhang stehen.

Die Verständlichkeit kann definitionsgemäß nur Werte zwischen 0 % und 100 % annehmen. Dadurch können unerwünschte Boden- und Deckeneffekte auftreten. CI-Träger, für welche der Test entwickelt wurde, unterscheiden sich sehr in ihrer Fähigkeit, Sprache zu verstehen. Aber auch bei einem einzelnen Patienten verändert sich nicht selten diese Fähigkeit mit der Tragedauer des CIs erheblich, so dass sich auch von daher die Ergebnisse nach Erstanpassung und etwa einem Jahr deutlich unterscheiden können. Aus diesem Grunde ermittelt man bei Sprachtests häufig die sogenannte Sprachverständlichkeitsschwelle.

Die Sprachverständlichkeitsschwelle kann auf zwei unterschiedliche Weisen bestimmt werden. Es kann der Sprachpegel bei konstantem Rauschpegel oder der Rauschpegel bei konstantem Sprachpegel einreguliert werden. Unter dem Aspekt einer den Alltag widerspiegelnden Situation wird für die erste Alternative angeführt, dass man in lauter Umgebung auch lauter spräche. Dieses Argument trifft nur teilweise zu. Normalhörende beginnen im Allgemeinen nicht lauter zu sprechen, wenn um sie herum mehrere Leute gleichzeitig sprechen. Hörbehinderte hingegen können in dieser Situation schon ernste Verständnisschwierigkeiten haben. Nur diese beginnen daher, lauter zu sprechen.

Da die Verständlichkeit mit wachsendem Sprachpegel aus zwei Gründen zunehmen kann, sind auf diese Weise gewonnene Ergebnisse schwierig zu interpretieren. Eine bessere Verständlichkeit kann nämlich erstens durch einen höheren Sprachpegel und zweitens durch einen günstigeren Signal-Rausch-Abstands erreicht werden. In dem Vortest wurde daher wie bei der Anwendung des HSM-Tests in Würzburg der Sprachpegel konstant gehalten und der Rauschpegel variiert. Darüber hinaus bildet diese Variante das Nichteingehen der Normalhörenden auf die Bedürfnisse der Hörbehinderten im Alltag realitätsnah ab.

Verschiedene Autoren haben speziell für Sprachtests Anforderungen formuliert, die bei der Realisierung des Tests erfüllt werden sollten (Ackermann 1969, Albrecht 1967, Versfeld 2000). So fanden die Aufnahmen in einer schallisolierten Kabine unter kontrollierten Bedingungen statt. Für die Aufsprache konnte ein professioneller Nachrichtensprecher gewonnen werden. Ein solcher-

maßen geschulter Sprecher ist darin geübt, Stimmlage und Lautstärke seiner Aussprache zu kontrollieren, sowie natürlich und flüssig zu sprechen. Seine Stimme klingt locker, ohne Verspannungen in der Hals- und Mundbodenmuskulatur, die Artikulation ist klar, aber nicht übertrieben. Nach Hahlbrock (1970) zeigen die Ergebnisse mit einem männlichen Sprecher eine kleinere Varianz als mit einem weiblichen.

Als Störgeräusch wurde sprachsimulierendes Rauschen benutzt, das aus der Aufsprache des Tests abgeleitet wurde. Damit wurde erreicht, dass das Rauschspektrum mit dem Langzeitspektrum der Sätze übereinstimmt. Die Testsätze werden dadurch in allen Frequenzbereichen gleichmäßig durch das Rauschen maskiert (Niemeyer 1976, Schultz-Coulon 1973, Hagerman 1982). Daraus resultiert eine steile Verständlichkeitsfunktion, welche dem Test eine hohe Diskriminationsfähigkeit verleiht.

Sätze und Rauschen wurden digital auf CD aufgenommen. Die digitale Aufzeichnung hat gegenüber der hergebrachten analogen Aufnahmetechnik mit Magnetbändern, Kassetten oder Schallplatten mehrere Vorteile: Die Qualität der Wiedergabe bleibt unverändert hoch, unabhängig wie oft der Datenträger abgespielt wird. Und es können ohne Qualitätsverlust beliebig viele Kopien angefertigt werden. Als weitere Vorteile bietet die CD einen größeren Signal-Rausch-Abstand, eine größere Dynamik, bessere Bedienbarkeit und einen schnelleren Zugriff auf bestimmte Wörter und Sätze. Die CD gestattet damit eine größtmögliche Flexibilität bei der Darbietung des Sprachmaterials (Kamm 1980, Alich 1985, von Wedel 1986, Moser 1987, Welzl-Müller 1989, Doerr 1994, Kießling 2000).

Bei den Verständlichkeitsmessungen wurde darauf geachtet, im erreichbaren Maß Standardisierbarkeit und Reproduzierbarkeit zu sichern. Die Darbietung erfolgte über hochwertige Lautsprecher. Um auch Einflüsse der Raumakustik zu kontrollieren, wurden die Tests in einer genormten Audiometrikammer durchgeführt. Außerdem kommt man mit der Darbietung über Lautsprecher der Alltagssituation näher als über Kopfhörer (Chung, Mack 1979). Vor allem aber ist der Test in dieser Weise mit CI-Trägern einfach durchzuführen. Die Probanden sahen den Lautsprecher mittig vor sich. Sie hörten die Listen

beidohrig, in der natürlichen Art des Hörens, ab. Manipulationen an den Ohren waren daher nicht notwendig. Schließlich führt auch beidohriges Hören zu einer steileren Verständlichkeitsfunktion (Ingold, Tschopp 1992).

Bei Satztests ist der Lerneffekt stärker ausgeprägt als bei anderen Materialien (Schultz-Coulon 1974). Erwartungsgemäß zeigen die Messungen einen signifikanten Lerneffekt von 4 % über die ersten zehn Listen. Offensichtlich vermögen sich die Probanden in den Test einzuhören. Rechnerisch ergibt sich daraus bei einer Steilheit von 10 %/dB eine Verbesserung in der Sprachverständlichkeitsschwelle von 0,4 dB. Das sind weniger als die von Wagener (2003a) angegebenen 2,2 dB für die dänische Variante des von Hagerman angegebenen und als unempfindlich eingeschätzten Testverfahrens.

Scheinbar registrieren einige Probanden, dass ihnen der Test mit der Zeit leichter fällt. Hierzu passt der Bericht zweier Probanden, die angaben, sie hätten das Gefühl, im Laufe der Untersuchung zunehmend besser verstanden zu haben. Bezogen auf die stets vorhandene Streuung in den Ergebnissen ($s_{\text{Liste}} = 6,8 \%$) erscheint der Übungseffekt aber tolerierbar.

Bei der Auswertung des Tests wurden dem Vorgehen beim HSM-Test folgend die richtig wiedergegebenen Wörter gezählt. Dies scheint die von den meisten Autoren akzeptierte Zählweise zu sein. Nicht notiert wurde, was die Probanden fälschlicherweise verstanden haben. Für jedes Wort wurde nur registriert, ob es richtig oder falsch verstanden wurde. Plomp und Mimpen werten in allen Anwendungen ihres (1979) publizierten Tests abweichend davon nur vollständig richtig verstandene Sätze aus.

Verständlichkeitsmessungen können mit einer Reihe von systematischen Fehlern behaftet sein. So konnten aus ökonomischen Gründen nicht alle Listen an allen Probanden gemessen werden. Probanden unterscheiden sich in ihrer Intelligenz, Aufmerksamkeit, Bereitschaft, nicht richtig gehörte Wörter durch sinnvoll erscheinende andere zu ersetzen, Belastbarkeit, Auffassungsgabe, Lernfähigkeit und vielen anderen Eigenschaften, die auf die Messung durchschlagen können (Bangert 1980). Gleichzeitig ist davon auszugehen, dass Aufmerksamkeit und Konzentration mit der Zeit nachlassen; die Probanden ermüden. Auch dies wirkt sich auf die Messergebnisse aus. Damit sich diese

Einflüsse im Mittel herausheben, begann jeder Proband die Untersuchung mit einer anderen Liste in zyklischer Reihenfolge. Dieses Vorgehen erlaubt später in der Phase der Auswertung, die dadurch verursachten Effekte zu kontrollieren und zu quantifizieren.

Bei Satztests lässt sich anders als bei Worttests eine vorgegebene Lautverteilung nur schwer einstellen, weil die einzelnen Wörter nicht so einfach ausgetauscht werden können (Niemeyer 1967). Trotzdem ist davon auszugehen, dass sich in dem vorgelegten Satztest ungefähr die Lautverteilung des Griechischen widerspiegelt, da die Sätze der gebräuchlichen Alltagssprache entnommen wurden. Die Auswahl und Zusammenstellung der Sätze zu den endgültigen Listen erfolgte zwar alleine auf Grund ihrer Verständlichkeit, aber da sich auch die ausgewählten Sätze wie die Sätze in der Urliste stark in ihrer Verständlichkeit unterscheiden, dürfte sich durch diese Nachauswahl die Lautverteilung nicht allzu viel verändert haben. Da die unterschiedlichen Laute einer Sprache unterschiedlich verständlich sind, die Listen aber auf gleiche Verständlichkeit abgeglichen wurden, so hat sich vermutlich auch die Lautverteilung der einzelnen Listen durch diese Nachauswahl eher angeglichen. Die Verständlichkeit der Listen beträgt 40 % bei einem Signal-Rausch-Abstand von -4 dB. Dies gilt für das benutzte Kollektiv normalhörender Griechen im Rahmen der stets vorhandenen statischen Schwankungen.

Die einzelnen Listen enthalten nur 10 Sätze. Wegen dieser geringen Anzahl sind die statistisch verursachten Abweichungen von der Lauthäufigkeitsverteilung in einer Liste größer als in der Gesamtstichprobe. Auch wirkt sich bei Hörbehinderten die Einschränkung der Verständlichkeit nicht gleichmäßig auf alle Laute aus, so dass sich bei der Untersuchung mit CI-Patienten größere Unterschiede in der Verständlichkeit zwischen den Listen, wie dies auch bei anderen Sprachtests der Fall ist, einstellen können.

Der hier vorgestellte griechische Satztest hat in der jetzigen Form 20 Listen mit je 10 Sätzen. Er gleicht damit der ursprünglichen Version des HSM-Tests, der auch aus Listen mit 10 Sätzen bestand. Falls man jeweils zwei Listen hintereinander abhört, so erreicht man Kompatibilität zur jetzigen Form des HSM-Tests mit 20 Listen. Zu diesem gibt es noch zwei weitere Unterschiede: a) die

Griechen pflegen die Dinge mehr als die Deutschen zu umschreiben. Um den gleichen Sachverhalt auszudrücken, bilden die Griechen längere Sätze als die Deutschen. b) Im HSM Test wird CCITT-Rauschen benutzt, für den griechischen Test wurde ein dem CCITT-Rauschen ähnliches Rauschen generiert, das exakt das Stimmspektrum des Sprechers nachbildet.

Nach Bortz et al. (1990) sollte jede Messung gewissen Kriterien genügen: Sie sollte objektiv, reliabel und valide sein.

Bei Verständlichkeitsmessungen muss man trotz aller technischen Fortschritte noch immer auf die einzigartige Fähigkeit des Menschen zurückgreifen, Sprache zu verstehen. Prinzipiell fließt dadurch ein, wenn auch im Vergleich mit den anderen Störquellen unbedeutendes, subjektives Element in die Messung ein. Eine Messung gilt nach Bortz als objektiv, wenn die Korrelation zwischen den Ergebnissen unterschiedlicher Untersucher den Wert von 0,9 überschreitet. In der vorliegenden Arbeit beurteilte ein einziger, normalhörender Versuchsleiter die Antworten aller Probanden, indem er jedes Wort als richtig oder falsch klassifizierte. Korrelationen zu anderen Untersuchern können folglich nicht angegeben werden. Es gibt aber keinen Hinweis, der ein unterschiedliches Verhalten des Versuchsleiters im Vergleich mit anderen Untersuchern vermuten ließe. Auf der anderen Seite wirkt sich die Beurteilung aller Probanden durch eine Person günstig auf die gemessene Varianz der Ergebnisse aus.

Eine Messung gilt als reliabel, wenn sie unter gleichen Bedingungen reproduzierbar ist. Da keine Test-Retest-Untersuchungen durchgeführt wurden, wird die Reliabilität anhand der Standardabweichung beurteilt. Der Ergebnisteil weist für jeden Satz die gefundene Standardabweichung aus. Betrachtet man das Verstehen eines Satzes als unabhängiges stochastisches Ereignis, so lässt sich daraus die Standardabweichung für die endgültigen Listen berechnen. Gemittelt über alle Listen ergibt sich ein Wert von ca. 6,8 % bei der eingestellten Verständlichkeit von 40 %. Dieser Wert liegt überraschend nahe bei dem Wert für eine Binomialverteilung. Für $n = 63$ Ereignisse, dem Wortumfang einer Liste, deren Wörter gleichwahrscheinlich mit einer Wahrscheinlichkeit von $p = 0,4$ verstanden werden, errechnet sich eine Standardabweichung von $s = 6,2$ %. Diese Übereinstimmung ist insofern überraschend,

da für ein Bernoulli Experiment zwei Voraussetzungen erfüllt sein müssen, die auf einen Sprachtest sicherlich nicht zutreffen. Daher zwingt die gute Übereinstimmung zwischen den Vorhersagen der Binomialverteilung und den experimentellen Daten zu der Annahme, der Unterschied wäre nur deshalb so gering, weil sich unterschiedliche Effekte gegenseitig kompensieren.

Definitionsgemäß müssen für eine Binomialverteilung zwei Voraussetzungen erfüllt sein. Erstens, die einzelnen Ereignisse in der Ereigniskette müssen mit der gleichen Wahrscheinlichkeit auftreten, und zweitens, sie müssen voneinander (stochastisch) unabhängig sein. Es widerspricht jeglicher Erfahrung, dass in einem Sprachtest, der als Ergebnis beispielsweise eine 50-prozentige Verständlichkeit ausweist, jedes Wort in diesem Test mit der Wahrscheinlichkeit von 50 % verstanden wird. Dazu sind Sprachsignale viel zu komplex und verschieden, als dass alle Wörter eines Tests gleichwahrscheinlich perzipiert würden (Büttner 1996).

Für das weitere soll an den Wert von 50% = 0,5 für die mittlere Verständlichkeit festgehalten werden, weil durch die dadurch vorgegebene Symmetrie die Diskussion erleichtert wird, ohne dass die Argumente an Allgemeingültigkeit verlören. Falls die Verständlichkeit eines Wortes mit der früher eingeführten Zufallsvariablen K beschrieben wird, dann ist die Varianz von K gegeben durch: $V(K) = p(1-p)$. Dies ist eine Parabel, die für $p = 0$ und $p = 1$ die Abszisse schneidet und für $p = 1/2$ ihr Maximum mit dem Wert $1/4$ annimmt. Gleichzeitig ist die Vertikale durch 0.5 die Symmetrieachse. Falls nun, wie das in der Realität der Fall ist, die Wörter mit einer von $p = 1/2$ verschiedenen Wahrscheinlichkeit wahrgenommen werden, dann ist für alle $p \neq 1/2$ das Produkt $p(1-p)$ kleiner als $1/4$. Falls weiterhin das Verstehen der Wörter in einem Satz als eine Abfolge von unabhängigen Ereignissen modelliert wird, so dass die Gesamtvarianz als Summe der Einzelvarianzen errechnet werden kann, so ist natürlich auch diese Summe kleiner als die n -malige Addition von $1/4$. Damit ergibt sich ein erster Effekt: Verglichen mit den Vorhersagen der Binomialverteilung führt eine breite Verteilung der Wahrscheinlichkeiten für die Verständlichkeit der einzelnen Wörter tendenziell zu einer kleineren Varianz der Sätze und damit auch zu einer kleineren Standardabweichung. Aus diesem

Grund kann, wie aus Abb. 7 zu ersehen ist, die Standardabweichung auch bei einem mittleren Satzverständnis extrem kleine Werte annehmen. Als Beispiel mag der Satz „O thios mu ine chimerinos kolimwitis“ mit einer Verständlichkeit von 51 % und einer Standardabweichung von 12 % dienen. Die ersten beiden Wörter in diesem Satz wurden immer und die letzten beiden Wörter niemals verstanden. Bei den beiden mittleren Wörtern liegt die Verständlichkeit bei 80 % und 30 %. Vier Wörter tragen also in diesem Fall überhaupt nicht zur Varianz bei und zwei nur eingeschränkt. Eine andere Konsequenz betrifft die Beobachtbarkeit. Da vier Wörter in die Kategorien „sicheres“ oder „unmögliches“ Ereignis fallen, reduzieren sich die beobachtbaren Konfigurationen in der Abfolge der verstandenen Wörter stark, nämlich auf $2^2 = 4$, so dass sie mit großer Wahrscheinlichkeit in einer Stichprobe von zehn Sätzen anzutreffen sind.

Falls man sich die Ergebnisse für die beiden mittleren Wörter in diesem Beispiel genauer ansieht, so drängt sich die Vermutung auf, dass die beiden Wörter keineswegs unabhängig voneinander verstanden werden. Bereits in den einführenden Abschnitten wurde darauf hingewiesen, dass sprachliche Äußerungen redundant sind. Die einzelnen Wörter sind aufeinander bezogen, und diese Verknüpfungen wirken wesentlich mit, wenn das Gehirn versucht, aus den verrauschten und zum Teil verdeckten Sprachsignalen ein sinnvolles Ganzes zu rekonstruieren. In einem entsprechenden stochastischen Modell sind deshalb statt der unbedingten nun bedingte Wahrscheinlichkeiten zu verwenden. Die Daten in dem genannten Beispiels suggerieren genau dies: In keinem Fall wurde das vierte Wort verstanden, wenn nicht vorher das dritte Wort richtig gehört wurde. Dagegen wurde das vierte Wort in 38 % der Fälle richtig verstanden, wenn vorher das dritte Wort richtig getroffen wurde. Diese Interaktion zwischen dem dritten und vierten Wort hat nun den umgekehrten Einfluss auf die Varianz; sie vergrößert diese. Die Gesamtvarianz errechnet sich nicht länger aus der Summe der Einzelvarianzen, sondern vergrößert sich um Zusatzglieder, die aus der Abhängigkeit der Ereignisse resultieren. Dieser Effekt ist durchaus nicht marginal. Bei zwei gleichverteilten, und unabhängigen Zufallsvariablen X und Y errechnet sich die Varianz von $X+Y$ als schlichte

Summe $\text{Var}(X) + \text{Var}(Y)$. Wegen der Gleichheit der Varianzen ergibt dies $2 \text{Var}(X)$ oder $2 \text{Var}(Y)$. Falls nun die beiden Variablen streng korreliert sind, d.h. Y nimmt immer den gleichen Wert wie X an, den ergibt $\text{Var}(X+Y) = \text{Var}(2 X) = 4 \text{Var}(X)$. Für die abhängigen Variablen hat sich die Varianz der Summe gegenüber den unabhängigen also verdoppelt. Für die 700 untersuchten Sätze vergrößert sich durch diesen Effekt die Standardabweichung im Mittel um 7,8 %. Zum Vergleich: Die mittlere Standardabweichung eines Satzes beträgt ungefähr 25%. Der Effekt ist also in den Daten deutlich sichtbar. Der beobachtete Zuwachs ist im Vorzeichen-Rang-Test nach Wilcoxon hoch signifikant ($p = 10^{-97}$).

Da die Verständlichkeit der Wörter teilweise miteinander korreliert, erscheint es nicht überraschend, dass mehrere Beispiele gefunden werden konnten, in denen die mittlere Verständlichkeit von 50 % dadurch zustande kommt, dass in der einen Hälfte nahezu jedes Wort und in der anderen keines verstanden wurde. Für eine Binomialverteilung wäre diese Wahrscheinlichkeit für eine Stichprobe aus 10 Sätzen so klein, dass man erwarten darf, keine einzige zu finden. Möglicherweise lag daher im Verstehen des ersten Wortes schon der Schlüssel zum Verstehen des ganzen restlichen Satzes.

Hagerman (1984) findet in seinen Untersuchungen eine kleinere Standardabweichung als die Binomialverteilung voraussagt. Er konnte diesen Unterschied auf ein akzeptables Maß verringern, indem er in der Formel für die Varianz die tatsächliche Anzahl der Wörter im Test durch eine fiktive, aus den Daten abgeleitete, ersetzte. Dadurch bewahrte er den Vorteil, welchen die aus der Binomialverteilung abgeleiteten Formeln bieten: Sie sind einfach anzuwenden und sie liefern eine Beziehung für die Abhängigkeit der Standardabweichung von der Verständlichkeit.

Zur Beurteilung der Validität des Tests fehlt ein Außenkriterium. Aber bislang sind keine Gründe erkennbar, warum der Test, die Fähigkeit einem alltäglichen Gespräch zu folgen, nur unzureichend abbilden sollte.

8 Zusammenfassung

In dieser Arbeit wurde nach dem Vorbild des deutschen HSM-Tests ein Satztest für das Griechische entwickelt. Speziell ist er wie der HSM-Test für die Untersuchung von Cochlea-Implantat-Trägern vorgesehen. Wie beim HSM-Test sollte das Testmaterial aus Alltagsätzen bestehen. Ein männlicher Rundfunksprecher sprach 914 griechische Sätze einer primären Sammlung mit möglichst ausgeglichener Lautstärke zweimal auf CD auf. Der in der Lautstärke homogenere Satz wurde für den Test ausgewählt. Es verblieben 711 Sätze, welche auf gleiche maximale Lautheit eingepegelt wurden. Aus den Sätzen wurde ein Rauschsignal abgeleitet, das in seiner spektralen Energieverteilung den Sätzen gleicht und in der Zeitstruktur weißem Rauschen. Für die vorliegende Untersuchung wurde eine neue CD gebrannt, welche auf der einen Stereospur 20 Listen zu je 35 Sätzen enthielt und auf der anderen Rauschen. Um sicherzustellen, dass zu Beginn des Sprachsignals die AGC des CIs eingeschwungen ist, setzt das Rauschsignal mit einer 200 ms langen Rampe schon 400 ms vor dem Sprachsignal ein.

In der Hauptstudie wurde die Verständlichkeit eines jeden Satzes gemessen. Dazu hörten 20 griechische, jugendliche Probanden jeweils 10 Listen bei einem Signal-Rausch-Abstand von -4 dB ab, insgesamt also 200 Listen. Da das Testmaterial aus 20 Listen bestand, wurden für jede Liste und damit auch für jeden Satz die Verständlichkeiten an 10 unterschiedlichen Probanden getestet. Mit der so gewonnenen Verständlichkeit der einzelnen Sätze wurden 20 neue Listen zu je 10 Sätzen erstellt, die gleich verständlich und von gleicher Struktur sind. Die Sätze wurden so zusammengestellt, dass die neuen Listen mit dem beschriebenen Probandenkollektiv eine Verständlichkeit von 40 % aufweisen. Die Listen setzen sich aus je 2 Frage- und 8 Aussagesätzen zusammen. Die Länge der Sätze variiert zwischen 4 und 9 Wörtern. Alle Listen enthalten an gleicher Position einen Satz gleicher Länge. Im Einzelnen haben die Sätze in der angegebenen Reihenfolge die Wortlängen: 9, 5?, 6?, 8, 7, 7, 6, 5, 4, 6 (? = Fragesatz). Jede Liste umfasst also 63 Wörter. Wie die Arbeit zeigt, konnten Ideen und Prinzipien, die dem deutschen HSM-Test zu Grund lagen, weitestgehend mit ähnlichen Ergebnissen auch im Griechischen verwirklicht werden.

9 Literaturverzeichnis

- Ackermann R (1969): Freiburger Sprachverständnistest und seine Apperzeption. HNO, 17: 89-91
- Albrecht R (1967): Bemerkungen zum Freiburger Sprachverständnistest. HNO 15: 77-80
- Alich G (1985): Anmerkungen zum Freiburger Sprachverständnistest (FST). Sprache-Stimme-Gehör 9: 1-6
- ANSI (1960) American National Standards Institute:
Method for measurement of monosyllabic word intelligibility.
American National Standards Institute, New York
- Bangert H (1980): Probleme bei der Ermittlung des Diskriminationsverlustes nach dem Freiburger Sprachtest. Audiol Akust 19: 166-170
- Barany R (1910): Worttabellen zur Hörprüfung.
Verh Dtsch Otol Ges, Dresden, 110-113
- Bertolli S, Probst R, Jordan P (1996): Das Hörhandicap – Eine Ergänzung zum audiometrischen Hörverlust. HNO 44: 376-384.
- Bezold F. (1896): Über den gegenwärtigen Stand der Hörprüfungen.
Z Ohrenheilkd 29: 1-28
- Bortz J, Lienert GA, Boehnke K (1990): Verteilungsfreie Methoden in der Biostatistik, Springer Berlin, Heidelberg, New York, London usw.
- Boothroyd A, Nitttrouer S (1988) Mathematical treatment of context effects in phoneme and word recognition. J Acoust Soc Am 84: 101-114
- Bosman A (1989): Speech perception by hearing impaired.
Doctoral thesis, Univ. Utrecht, The Netherlands

- Bosman A (1992): Review of speech audiometric tests. In Kollmeier (Hrsg.)
Moderne Verfahren der Sprachaudiometrie, 11-34,
Audiologische Akustik, Bd. 1, Median, Heidelberg
- Brand T, Kollmeier B (2002): Vorhersage der Sprachverständlichkeit in Ruhe
und im Störgeräusch aufgrund des Reintonaudiogramms.
3. DAGA Jahrestagung, CD-ROM
- Brinkmann K (1974): Die Neuaufnahme der „Wörter für Gehörprüfung mit
Sprache“. Z Hörgeräte-Akustik 13: 12-40
- Bronkhorst AW, Bosmann AJ, Smoorenburg GF (1993): A model for context
effects in speech recognition. J Acoust Soc Am 93: 499-509
- Bronkhorst AW, Brand T, Wagener K (2002): Evaluation of context effects in
sentences recognition. J Acoust Soc Am 111: 2874-86
- Büttner, J (1996): Sprachaudiometrische Untersuchungen an Einzelwörtern.
Diss. Universität Würzburg
- Chung DY, Mack B (1979): The effect of masking by noise on word
discrimination scores in listeners with normal hearing and with noise-
induced hearing loss. Scand Audiol 8: 139-143
- Dillier N, Spillmann T (1992): Deutsche Version der Minimal Auditory Capability
(MAC)-Test-Batterie, Anwendungen bei Hörgeräte- und CI-Trägern
mit und ohne Störlärm. In Kollmeier (Hrsg.) Moderne Verfahren der
Sprachaudiometrie, 238-263, Audiologische Akustik, Bd. 1, Median,
Heidelberg
- DIN 45621: Sprache für Gehörprüfung – Teil 1: Ein- und mehrsilbige Wörter
Beuth Verlag, Ausgabe : 1995-08
- Dölp U (1992): Entwicklung eines offenen Logatomtests zur Beurteilung des
Sprachgehörverlustes – erste Ergebnisse einer Exemplarischen
Untersuchung. In: Kollmeier B. (Hrsg.), Moderne Verfahren der

Sprachaudiometrie, 169-195, Audiologische Akustik, Bd. 1,
Median, Heidelberg

Döring WH, Hamacher V (1992): Neue Sprachverständlichkeitstests in der
Klinik: Aachener Logatomtest und „Dreinsilber“ Test mit Störschall.
In: Kollmeier B. (Hrsg.), Moderne Verfahren der Sprachaudiometrie,
137-168, Audiologische Akustik, Bd. 1, Median, Heidelberg

Doerr L (1994): Vergleichende sprachaudiometrische Untersuchungen mit
Kompaktdiskette (CD) und Tonbandkassette. HNO 42: 493-498.

Duden (1977): Das große Wörterbuch der deutschen Sprache in sechs Bänden.
Band 2, Bibliographisches Institut, Mannheim, Wien, Zürich

Egan JP (1948): Articulation testing methods. Laryngoscope 58: 955-991

Egan JP (1957): Remarks on rare PB words. J Acoust Soc Am 29: 751

Fastl H (1987): Ein Störgeräusch für die Sprachaudiometrie.
Audiol Akust 26: 2-13

Fastl H, Zwicker E, Fleischer R (1990): Beurteilung der Verbesserung der
Sprachverständlichkeit in einem Hörsaal. Acustica 71: 287-292

Feldmann H (1960): Die geschichtliche Entwicklung der Hörprüfungsmethoden.
Thieme, Stuttgart

Feldmann H (1965): Dichotischer Diskriminationstest, eine neue Methode zur
Diagnostik zentraler Hörstörungen.
Arch Ohr Nas Kehlk Heilkd 184: 294-329

Feldmann H (2004): 200 Jahre Hörprüfungen mit Sprache, 50 Jahre deutsche
Sprachaudiometrie – ein Rückblick.
Laryngol Rhinol Otol 83: 735-742

Fletcher H (1953): Speech and hearing in communication. D. van Nostrand
Company, New York, Toronto, London

- Fodor JA (1983): The modularity of mind. MIT Press, Cambridge/MA
- Gavrilidou M, Labropoulou P, Mantzani E, Roussou S (1999): Specifications for a computational morphological lexicon of modern Greek. In A. Mozer (ed.), Greek Linguistics 97, Proceedings of the 3rd International Conference on the Greek Language. Athens: Ellinika Grammata, pp. 929-936.
- Gavrilidou M (2002): The Hellenic National Corpus on-line. Revue Belge de Philologie et d' Histoire 80/2002, Bruxelles, Société pour le progrès des études philologiques et historiques, 1003-1015.
- Hagerman B (1982): Sentences for testing speech Intelligibility in noise. Scand Audiol 11: 79-87
- Hagerman B (1984): Some aspects of methodology in speech audiometry. Scand Audiol Suppl 21: 3-25
- Hagerman B, Kinnefors K (1995): Efficient adaptive methods for measuring speech recognition threshold in quiet and in noise. Scand Audiol 24: 71-7
- Hahlbrock KH (1953): Über Sprachaudiometrie und neue Wörtertteste. Archiv Ohr- usw. Heilkd u. Z Hals- usw. Heilkd.162: 394-431
- Hahlbrock KH (1970): Sprachaudiometrie, Grundlagen und praktische Anwendung einer Sprachaudiometrie für das deutsche Sprachgebiet. 2. Auflage, Thieme Verlag
- Hatzigeorgiu N., Gavrilidou M., Piperidis S., Karayannis G., Papakostopoulou A., Spiliotopoulou A., Vakalopoulou A., Labropoulou P., Mantzari E., Papageorgiou H., Demiros I (2000): Design and implementation of the online ILSP Greek Corpus. Proceedings of LREC 2000, Athen, Griechenland, <http://www.sdjt.si/bib/lrec00/ps/336.ps>

- Heller O (1992): Frequenzspezifische Sprachaudiometrie. In Kollmeier (Hrsg.)
 Moderne Verfahren der Sprachaudiometrie, 264-277,
 Audiologische Akustik, Bd. 1, Median, Heidelberg
- Helmholtz HLF von (1863): Die Lehre von den Tonempfindungen als
 physiologische Grundlage für die Theorie der Musik.
 Vieweg, Braunschweig
- Hirsh IJ, Davis H, Silverman SR, Eldert E, Benson RW (1952): Development of
 materials for speech audiometry. J Speech Disord 17: 321-337
- Hochmair-Desoyer I, Schulz E., Moser L, Schmidt M (1997): The HSM sentence
 test as a tool for evaluating the speech understanding in noise of
 cochlear implant users. Am J Otol 18: No. 6 (Suppl), 83
- Iliadou V, Fourakis M, Vakalos A, Hawks JW, Kaprinis G (2006): Bi-syllabic,
 modern Greek word lists for use in word recognition tests.
 Int J Audiol 45: 74-82
- Ingold L, Tshopp K (1992): Praktische Aspekte zur Sprachaudiometrie mit
 Störgeräusch. Sprache-Stimme-Gehör 16: 116-120.
- Itard JMG (1821) Traité des maladies de l'oreille et de l'audition. Paris
- Kalikow DN, Stevens KN (1977): Development of a test of speech intelligibility
 in noise using sentence materials with controlled word predictability.
 J Acoust Soc Am 61: 1337-1351.
- Kamm C (1980): Use of digitized speech materials in audiological research.
 J Speech Hear Res 23: 709-721
- Kastellis G (2009): Persönliche Mitteilung
- Kastellis G, Battmer RD (1978): Sprachtest für Griechen in Deutschland.
 Arch Ohr Nas Kehlk Heilkd 219: 420-421

- Kießling J, Schubert M, Wagner I (1994): Sprachverständlichkeitsmessungen an Normalhörenden und Schallempfindungsschwerhörigen – fünf Sprachtests im Vergleich. Teil 1 und 2.
Audiol Akust, Vol. 33: 6-15, 11-15
- Kießling J (1996): Hörgeräteauswahl und Anpassung. In Lehnhardt, Praxis der Audiometrie, Thieme, Stuttgart, New York
- Kießling J (2000): Moderne Verfahren der Sprachaudiometrie.
Laryngol Rhinol Otol 79: 633-635
- Kogias A (1961): How to choose words for speech audiometry.
Acad Med 25:265
- Kogias A (1974): Oto-Rhino-Laryngologie: Klinische Untersuchung, diagnostische Methodologie und interessante Themen. S. 32-39,
Gr. Parisianos Verlag, Athen 1974.
- Kollmeier B (1992): Weiterentwicklung des Reimtests nach Sotscheck. In:
Kollmeier B. (Hrsg.), Moderne Verfahren der Sprachaudiometrie,
216-237, Audiologische Akustik, Bd. 1, Median, Heidelberg
- Kollmeier B, Wesselkamp M (1997): Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment. J Acoust Soc Am 102: 2412-2421
- Kramer W (1861): Die Ohrenheilkunde der Gegenwart (1860).
Hirschwald, Berlin
- Lampert H (1923): Die Mängel der bisherigen Hörprüfung mit der Sprache.
Z Hals-Nasen-Ohrenheilkd 5: 69-91
- Langenbeck B (1963): Leitfaden der praktischen Audiometrie.
Thieme Verlag, Stuttgart

- Lehiste I, Peterson GE (1959): Linguistic considerations in the study of speech intelligibility. *J Acoust Soc Amer* 31: 280-286
- Lehnhardt E, Janssen THJ, Kiessling J (1996): *Praxis der Audiometrie*. 7. Auflage, Thieme, Stuttgart
- Lyregaard PE (1982): Frequency-selectivity and speech-intelligibility in noise. *Scand Audiol Suppl* 15: 113-122
- Manolidis L (1964): Development and use of speech audiometry in the Greek language. Diss. Aristoteles Universität Thessaloniki, Griechenland
- Marslen-Wilson W, Tyler LK (1980): The temporal structure of spoken language understanding. *Cognition* 8: 1-71
- Marslen-Wilson WD (1987): Functional parallelism in spoken word-recognition. *Cognition* 25: 71-102
- Martin M (1997): *Speech Audiometry*. 2nd Edition, Whurr Publishers Ltd
- Matzker J (1958): Ein binauraler Hörsynthese-Test zum Nachweis zerebraler Hörstörungen. Thieme, Stuttgart
- McClelland JL, Elman JL (1986): The TRACE model of speech perception. *Cogn Psychol* 18: 1-86
- Mikros G, Hatzigeorgiu N, Carayannis G (2005): Basic quantitative characteristics of the modern Greek language using the Hellenic national corpus. *J Quant Linguist* 12: 167-184
- Miller G.A., G.A. Heise, W. Lichten (1951): The intelligibility of speech as a function of the context of the test material. *J Exp Psychol* 41 329-335
- Morton J (1969): The interaction of information in word recognition. *Psychol Rev* 76: 165-178

- Moser LM (1987): Das Würzburger Hörfeld, ein Test für prothetische Audiometrie. HNO 35: 318-321
- Neuss O (1982): Beitrag zu den Beziehungen zwischen Ton- und Sprachaudiometrie. Arch Otorhinolaryngol 234: 277-283
- Niemeyer W, Beckmann G (1962): Ein sprachaudiometrischer Satztest. Arch. Ohr-, Nasen-, u. Kehlk-Heilkd 180: (Kongressbericht), 742-749
- Niemeyer W (1967): Sprachaudiometrie mit Sätzen. HNO 15: 335-343
- Niemeyer W (1976): Speech audiometry and fitting of hearing-aids in noise. Audiology 15: 421-427
- Niklès JM, Tschopp K (1996): Audiologische Grundlagen des Basler Satztests. Audiol Akust 2: 70-75
- OLSA, Oldenburger Satztest (2000): Handbuch und Hintergrundwissen. HörTech gGmbH, Marie-Curie-Straße 2, D-26129 Oldenburg
- Papafragou K (1996): Audiologie. Parisianos G. Athen
- Pekkarinen E, Salmivalli A, Suonpää J (1990): Effects of noise on word discrimination by subjects with impaired hearing, compared with those with normal hearing. Scand Audiol 19: 31-36
- Pfingsten GW (1804) Gehörmesser zur Untersuchung der Gehörfähigkeit galvanisierter Taubstummer, in besonderer Rücksicht auf die Erlernung der artikulierten Tonsprache und deren Elemente gegründet. Kiel
- Plomp R., Mimpen A.M. (1979): Improving the reliability of testing the speech reception threshold for sentences. Audiology 18: 43-52
- Reuter C (1904): Beitrag zur Prüfung der Gehörschärfe mit der Flüsterstimme. Z Ohrenheilkd 47: 91-99

Schmaltz E (1846): Erfahrungen über die Krankheiten des Gehörs und ihrer Heilungen. Teubner, Leipzig

Schmidt M, Hochmair-Desoyer I, Schulz E, Moser L (1997): Der HSM-Satztest. Fortschritte der Akustik, Kolloquien, Plenarvorträge und Fachbeiträge der DAGA 97, 93-94

Schorn K, Zwicker E (1986): Klinische Untersuchungen zum Zeitauflösungsvermögen des Gehörs bei verschiedenen Hörschädigungen. Audiol Akust 25, 170-184

Schubert K (1952): Über einen Sprachprüftest für Schwerhörigengeräte. Archiv Ohr- Heilkd Z Hals- Heilkd 162: 67-84

Schultz-Coulon HJ (1973): Über die Bedeutung des Umweltgeräusches für den Hochtonschwerhörigen. HNO 21: 26-32

Schultz-Coulon HJ (1974): Sprachaudiometrie mit Sätzen und Geräusch. Laryng Rhinol 53: 734-749.

Shannon CE (1951): Prediction and entropy of printed English. Bell Syst Tech J 30: 30-84

Smooenburg GF, de Laat JAPM, Plomp R (1982): The effect of noise-induced hearing loss on the intelligibility of speech in noise. Scand Audiol Suppl 16: 123-133

Smooenburg GF (1992): Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram. J Acoust Soc Am 91: 421-437

Sotscheck J (1976a): Methoden zur Messung der Sprachgüte I: Verfahren zur Bestimmung der Satz- und der Wortverständlichkeit. Fernmelde-Ingenieur, 30(10): 1-31

Sotscheck J (1976b): Methoden zur Messung der Sprachgüte II.: Verfahren zur Bestimmung der Satz- und der Wortverständlichkeit. Fernmelde-Ingenieur, 30(12): 1-32

Sotscheck J (1985): Sprachverständlichkeit bei additiven Störungen. Acustica 57: 257-267

Stumpf C (1926): Die Sprachlaute. Berlin

Sukowski H, Brand T, Wagnener KC, Kollmeier B (2009): Untersuchung zur Vergleichbarkeit des Freiburger Sprachtests mit dem Göttinger Satztest und dem Einsilber-Reimtest nach Wallenberg und Kollmeier. HNO 57: 239-250

Trimmis N, Papadeas E, Papadas T, Naxakis S, Papathanasopoulos P, Goumas P (2006): Speech audiometry: The development of modern Greek word lists for suprathreshold word recognition testing. Mediterr J Otol 3: 117-126

Trimmis N, Markatos N, Malaperdas K, Papadas T (2007): Development of an Audio Compact Disc for Speech Audiometry Testing. 8th EFAS Congress / 10th Congress of the German Society of Audiology [http://www.uzh.ch/orl/dga2007/program/scientificprogram/Trimmis N. Markatos N. Malaperdas K. Papadas T. 2 .pdf](http://www.uzh.ch/orl/dga2007/program/scientificprogram/Trimmis_N._Markatos_N._Malaperdas_K._Papadas_T._2_.pdf)

Tröltsch AF von (1867): Lehrbuch der Ohrenheilkunde mit Einschluss der Anatomie des Ohres. Stahel, Würzburg

Tschopp K, Probst R (1992): Aktuelle Fragestellungen und Entwicklungen in der Sprachaudiometrie. Otorhinolaryngol Nova 2: 306-312

Versfeld NJ, Daalder L, Festen JM (2000): Method for the selection of sentence materials for efficient measurement of the speech reception threshold. J Acoust Soc Am 107: 1671-1684

- Wagener K, Kühnel V, Kollmeier B (1999a): Entwicklung und Evaluation eines Satztests für die deutsche Sprache: Teil I: Design des Oldenburger Satztests. Z Audiol 38: 4-15
- Wagener K, Brand T, Kollmeier B (1999b): Entwicklung und Evaluation eines Satztests für die deutsche Sprache: Teil II: Optimierung des Oldenburger Satztests. Z Audiol 38: 44-56
- Wagener K, Brand T, Kollmeier B (1999c): Entwicklung und Evaluation eines Satztests für die deutsche Sprache: Teil III: Evaluation des Oldenburger Satztests. Z Audiol 38: 86-95
- Wagener K (2003a): Factors influencing sentence intelligibility in noise. Diss. Universität Oldenburg
- Wagener K, Josvassen JL, Ardenkjær R (2003b): Design, optimization and evaluation of a Danish sentence test in noise. Internat. J Audiol 42: 10-17.
- Wedel H von (1986): Entsprechen die sprachaudiometrischen Untersuchungsverfahren den heutigen Anforderungen in Klinik und Praxis? HNO 34: 71-74
- Wedel H von (1985): Untersuchungen zur Sprachdiskrimination bei umweltspezifischen Störgeräuschen. Laryngol Rhinol Otol 64: 430-435
- Welge-Lüssen A, Hauser R, Erdmann J, Schwob C, Probst R (1997): Sprachaudiometrie mit Logatomen. Laryngol Rhinol Otol 76: 57-64
- Welzl-Müller K, Graber A, Stephan K (1989): Sprachaudiometrie: Vergleich Tonband – Compact Disk. Audiol Akust 3: 96-107
- Werra P von, Tschopp K, Schneider R (1995): Einfluss der linguistischen Kompetenz auf die sprachaudiometrischen Ergebnisse mit dem Basler Satztest. HNO 43: 98-103

Wesselkamp M, Kliem K, Kollmeier B (1992): Erstellung eines optimierten Satztestes in deutscher Sprache. In Kollmeier (Hrsg.) Moderne Verfahren der Sprachaudiometrie, 330-343, Audiologische Akustik, Bd. 1, Median, Heidelberg

Wolf O (1899): Die Hörprüfung mittels der Sprache. Z. Ohrenheilk. 34: 289-311

Züst HJ, Tschopp K (1995): Der Kontexteffekt bei der Spracherkennung von Sätzen. Laryngol Rhinol Otol 74: 259-263

Zwaardemaker H, Quix FH (1905): Onderzoegenkingen van het phys. Laboratorium te Utrechtse Hoogeschool. Int Zbl Ohrenheilkd 3: 233-240

Zwicker E, Fastl H (1999): Psychoacoustics – Facts and Models. Springer, Heidelberg, New York

10 Tabellarischer Anhang

10.1 Verständlichkeit der Listen

| | Liste 1 | Liste 2 | Liste 3 | Liste 4 | Liste 5 | Liste 6 | Liste 7 | Liste 8 | Liste 9 | Liste 10 | Liste 11 | Liste 12 | Liste 13 | Liste 14 | Liste 15 | Liste 16 | Liste 17 | Liste 18 | Liste 19 | Liste 20 |
|-----|---------|---------|---------|---------|---------|---------|---------|---------|---------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| P1 | 21 | 24 | 18 | 29 | 19 | 22 | 23 | 23 | 18 | 29 | | | | | | | | | | |
| P2 | | 28 | 34 | 44 | 28 | 24 | 37 | 55 | 39 | 46 | 54 | | | | | | | | | |
| P3 | | | 26 | 21 | 31 | 33 | 37 | 27 | 31 | 30 | 49 | 55 | | | | | | | | |
| P4 | | | | 23 | 24 | 35 | 25 | 34 | 33 | 41 | 46 | 60 | 44 | | | | | | | |
| P5 | | | | | 34 | 31 | 42 | 44 | 33 | 40 | 47 | 58 | 49 | 46 | | | | | | |
| P6 | | | | | | 34 | 35 | 40 | 30 | 32 | 46 | 66 | 37 | 50 | 42 | | | | | |
| P7 | | | | | | | 22 | 27 | 31 | 29 | 39 | 45 | 28 | 23 | 27 | 35 | | | | |
| P8 | | | | | | | | 29 | 27 | 36 | 40 | 53 | 33 | 40 | 32 | 44 | 49 | | | |
| P9 | | | | | | | | | 23 | 34 | 36 | 54 | 41 | 40 | 39 | 49 | 54 | 55 | | |
| P10 | | | | | | | | | | 32 | 46 | 56 | 34 | 46 | 37 | 38 | 44 | 49 | 53 | |
| P11 | | | | | | | | | | | 40 | 54 | 42 | 48 | 37 | 45 | 43 | 48 | 43 | 38 |
| P12 | 21 | | | | | | | | | | | 52 | 37 | 42 | 32 | 47 | 51 | 50 | 49 | 31 |
| P13 | 25 | 20 | | | | | | | | | | | 41 | 44 | 33 | 49 | 54 | 43 | 50 | 36 |
| P14 | 19 | 28 | 19 | | | | | | | | | | | 31 | 22 | 41 | 33 | 33 | 39 | 25 |
| P15 | 29 | 13 | 16 | 16 | | | | | | | | | | | 26 | 31 | 37 | 31 | 37 | 24 |
| P16 | 31 | 27 | 16 | 30 | 25 | | | | | | | | | | | 39 | 36 | 37 | 50 | 30 |
| P17 | 21 | 14 | 15 | 28 | 22 | 27 | | | | | | | | | | | 38 | 24 | 33 | 30 |
| P18 | 27 | 21 | 19 | 16 | 24 | 22 | 29 | | | | | | | | | | | 40 | 42 | 29 |
| P19 | 24 | 14 | 19 | 19 | 15 | 18 | 25 | 23 | | | | | | | | | | | 36 | 21 |
| P20 | 28 | 24 | 28 | 28 | 17 | 29 | 37 | 31 | 28 | | | | | | | | | | | 29 |

Tabelle 4: Verständlichkeit der Listen in Prozent der Wörter aufgeschlüsselt nach Probanden

10.2 Satzlisten

Liste 1: (Wortverständlichkeit V = 40 % ± 6,4 %)

- 1 To spiti mas echi dio patomata.
- 2 Fewgume gia tin ekdromi awrio to proi.
- 3 Pion tha psifisete stis ekloges?
- 4 Ta kalokeria mas aresun i diakopes konta sti thalassa.
- 5 Grafi panta me penna.
- 6 Meta apo polli skepsi sinenese sto schedio mas.
- 7 Posi zachari thelete ston kafe sas?
- 8 To riaki schimatizi pollus katarraktes.
- 9 Woithise tin giagia na perasi ton dromo.
- 10 I gnosi xenon glosson ine chrisimi.

Κατάλογος 1:

- 1 Το σπίτι μας έχει δύο πατώματα.
- 2 Φεύγουμε για την εκδρομή αύριο το πρωί.
- 3 Ποιόν θα ψηφίσετε στις εκλογές;
- 4 Τα καλοκαίρια μας αρέσουν οι διακοπές κοντά στη θάλασσα.
- 5 Γράφει πάντα με πένα.
- 6 Μετά από πολλή σκέψη συναίνεσε στο σχέδιό μας.
- 7 Πόση ζάχαρη θέλετε στον καφέ σας;
- 8 Το ρυάκι σχηματίζει πολλούς καταρράκτες.
- 9 Βοήθησε την γιαγιά να περάσει τον δρόμο.
- 10 Η γνώση ξένων γλωσσών είναι χρήσιμη.

Liste 2: (Wortverständlichkeit V = 40 % ± 7,7%)

- 1 Diawazi efcharista asti-nomika mith-istorimata.
- 2 Meleta me pathos tin archea elliniki istoria.
- 3 Ti ora xipnate to proi?
- 4 Ta pedia tis polis mathenun ta zoa apo fotografies.
- 5 Paragilame idi to epidorpio .
- 6 I agora itan gemati kosmo logo ton ekptoseon.
- 7 Poso akriwine to isitirio tu trenu?
- 8 Sto spiti ine panta kalitera.
- 9 Anixame to kenurgio wareli ke sas proskalume.
- 10 I istoria su diskola ginete pistefti.

Κατάλογος 2:

- 1 Διαβάζει ευχάριστα αστυ-νομικά μυθ-ιστορήματα.
- 2 Μελετά με πάθος την αρχαία ελληνική ιστορία.
- 3 Τι ώρα ξυπνάτε το πρωί;
- 4 Τα παιδιά της πόλης μαθαίνουν τα ζώα από φωτογραφίες.
- 5 Παραγγείλαμε ήδη το επιδόρπιο.
- 6 Η αγορά ήταν γεμάτη κόσμο λόγω των εκπτώσεων.
- 7 Πόσο ακρίβηνε το εισιτήριο του τραίνου;
- 8 Στο σπίτι είναι πάντα καλύτερα.
- 9 Ανοίξαμε το καινούργιο βαρέλι και σας προσκαλούμε.
- 10 Η ιστορία σου δύσκολα γίνεται πιστευτή.

Liste 3: (Wortverständlichkeit V = 40 % ± 6,3 %)

- 1 Ewrexe poli ke I dromi plimirisan.
- 2 Ta elafia trechun tromagmena apo ton pirobolismo.
- 3 Mipos sinantises ton anipsio mu ?
- 4 O markadoros briskete pano sto trapezi, dipla sto tetradio.
- 5 Wgike protathlitis sto skaki.
- 6 Itan anetimi na antimetvpsun tin chthesini chiono-thiella.
- 7 Sas aresi o peripatos stin exochi?
- 8 Sinithizi na fora fardia rucha.
- 9 Tha prospathiso na ton wro sto tilefono.
- 10 I efimerides grafun ke polla psemmata.

Κατάλογος 3:

- 1 Έβρεξε πολύ και οι δρόμοι πλημμύρισαν.
- 2 Τα ελάφια τρέχουν τρομαγμένα από τον πυροβολισμό.
- 3 Μήπως συνάντησες τον ανηψιό μου;
- 4 Ο μαρκαδόρος βρίσκεται πάνω στο τραπέζι, δίπλα στο τετράδιο.
- 5 Βγήκε πρωταθλητής στο σκάκι.
- 6 Ήταν ανέτοιμοι να αντιμετωπίσουν την χθεσινή χιονο-θύελλα.
- 7 Σας αρέσει ο περίπατος στην εξοχή;
- 8 Συνηθίζει να φορά φαρδιά ρούχα.
- 9 Θα προσπαθήσω να τον βρω στο τηλέφωνο.
- 10 Οι εφημερίδες γράφουν και πολλά ψέματα.

Liste 4: (Wortverständlichkeit V = 40 % ± 7,3 %)

- 1 M' avta pu kanis ginese antipathis.
- 2 Tis aresi na pezi me tis kukles.
- 3 Pote tha pame gia ski?
- 4 To aero-plano efige me trofima gia tin Afriki.
- 5 Eftiaxa keik me fruta.
- 6 I enallagi ton kerikon fenomenon ine axio-simioti.
- 7 Poso apechi apo edo to chorio?
- 8 Sto periwoli tus kalliergun marulia.
- 9 I prosklisi ischii mono gia to fagito.
- 10 Akui ta nea apo to radiofono.

Κατάλογος 4:

- 1 Μ' αυτά που κάνεις γίνεσαι αντιπαθής.
- 2 Της αρέσει να παίζει με τις κούκλες.
- 3 Πότε θα πάμε για σκι;
- 4 Το αερο-πλάνο έφυγε με τρόφιμα για την Αφρική.
- 5 Έφτιαξα κέικ με φρούτα.
- 6 Η εναλλαγή των καιρικών φαινομένων είναι αξιο-σημείωτη.
- 7 Πόσο απέχει από εδώ το χωριό;
- 8 Στο περιβόλι τους καλλιεργούν μαρούλια.
- 9 Η πρόσκληση ισχύει μόνο για το φαγητό.
- 10 Ακούει τα νέα από το ραδιόφωνο.

Liste 5: (Wortverständlichkeit V = 40 % ± 6,2 %)

- 1 Ta Christugenna stilame polles efchetiries kartes.
- 2 Tin Megali Ewdomada pigeni kathimerina stin ekklesia.
- 3 Iparchi edo konta kapio estiatorio?
- 4 I ergates den plirononte kala ke katewenun se apergia.
- 5 Odigi mia palia sakaraka.
- 6 Tha borusame na taxidepsume an iche kalo kero.
- 7 Poso chronon ine o pateras su?
- 8 Sta pedia aresi to pegnidi.
- 9 I mitera prospathi na parigorisi to pedi.
- 10 Polis kosmos sigentrothike stin ekthesi zografikis.

Κατάλογος 5:

- 1 Τα Χριστούγεννα στείλαμε πολλές ευχετήριες κάρτες.
- 2 Την Μεγάλη Εβδομάδα πηγαίνει καθημερινά στην εκκλησία.
- 3 Υπάρχει εδώ κοντά κάποιο εστιατόριο;
- 4 Οι εργάτες δεν πληρώνονται καλά και κατεβαίνουν σε απεργία.
- 5 Οδηγεί μια παλιά σακαράκα.
- 6 Θα μπορούσαμε να ταξιδέψουμε αν είχε καλό καιρό.
- 7 Πόσο χρονών είναι ο πατέρας σου;
- 8 Στα παιδιά αρέσει το παιγνίδι.
- 9 Η μητέρα προσπαθεί να παρηγορήσει το παιδί.
- 10 Πολύς κόσμος συγκεντρώθηκε στην έκθεση ζωγραφικής.

Liste 6: (Wortverständlichkeit V = 40 % ± 5,7 %)

- 1 O skilos kimate sto spitaki tu.
- 2 To fidi tiligete giro apo to kladi.
- 3 Elenxes to aftokinito prin figis?
- 4 Mu fenete oti afto to katastima echi akriwes times.
- 5 Apopse echi nichterini wardia.
- 6 Milai poli choris na lei tipote to usiastiko.
- 7 Su aresi to kenurgio mu kurema?
- 8 Akume to treno na fthani.
- 9 Fisika ke tha ntithume maskarades tis Apokries.
- 10 To Sawato ichame giorti sti spiti.

Κατάλογος 6:

- 1 Ο σκύλος κοιμάται στο σπιτάκι του.
- 2 Το φίδι τυλίγεται γύρω από το κλαδί.
- 3 Έλεγξες το αυτοκίνητο πριν φύγεις;
- 4 Μου φαίνεται ότι αυτό το κατάστημα έχει ακριβές τιμές.
- 5 Απόψε έχει νυχτερινή βάρδια.
- 6 Μιλάει πολύ χωρίς να λέει τίποτε το ουσιαστικό.
- 7 Σου αρέσει το καινούργιο μου κούρεμα;
- 8 Ακούμε το τραίνο να φθάνει.
- 9 Φυσικά και θα ντυθούμε μασκαράδες τις Απόκριες.
- 10 Το Σάββατο είχαμε γιορτή στο σπίτι.

Liste 7: (Wortverständlichkeit V = 39,7 % ± 6,6 %)

- 1 O pagotatzis dialali tin pramatia tu.
- 2 Girise me to podilato oli tin Ellada.
- 3 Thelete na agorasete ena sali?
- 4 O kipos fenete thavmasios tora pu anthisan ta luludia.
- 5 Nomizo pos erchete kategida.
- 6 Se parakalo, mazedse ta fagita apo to trapezi.
- 7 Xeris to paramithi tu Konto-rewithuli?
- 8 I thalassa edo watheni apotoma.
- 9 Fetos to kalokeri tha kani polli zesti.
- 10 I kuzina ine aristera ston diadromo.

Κατάλογος 7:

- 1 Ο παγωτατζής διαλαλεί την πραμάτεια του.
- 2 Γύρισε με το ποδήλατο όλη την Ελλάδα.
- 3 Θέλετε να αγοράσετε ένα σάλι;
- 4 Ο κήπος φαίνεται θαυμάσιος τώρα που άνθισαν τα λουλούδια.
- 5 Νομίζω πως έρχεται καταιγίδα.
- 6 Σε παρακαλώ, μάζεψε τα φαγητά από το τραπέζι.
- 7 Ξέρεις το παραμύθι του Κοντο-ρεβουθούλη;
- 8 Η θάλασσα εδώ βαθαίνει απότομα.
- 9 Φέτος το καλοκαίρι θα κάνει πολλή ζέστη.
- 10 Η κουζίνα είναι αριστερά στο διάδρομο.

Liste 8: (Wortverständlichkeit V = 40 % ± 7,6 %)

- 1 Prosochi, giati to riaki ine wathi.
- 2 Kriologise ke epese sto krewati me pireto.
- 3 Itan kurastiko to taxidi sas?
- 4 Otan erthis fere ke ti fotografiki michani mazi su.
- 5 Odigise prosektika stin diastawrosi.
- 6 Ta magazia nichterinis diaskedasis xefitronun san ta manitaria.
- 7 Lipete poli kero apo to spiti?
- 8 Ine poli efcharistos gia pareia.
- 9 Elpizo na ise sinepis sto rantewu mas.
- 10 Mi kanis polli fasaria sto mathima.

Κατάλογος 8:

- 1 Προσοχή, γιατί το ρυάκι είναι βαθύ.
- 2 Κρυολόγησε και έπεσε στο κρεβάτι με πυρετό.
- 3 Ήταν κουραστικό το ταξίδι σας;
- 4 Όταν έρθεις φέρε και τη φωτογραφική μηχανή μαζί σου.
- 5 Οδήγησε προσεκτικά στη διασταύρωση.
- 6 Τα μαγαζιά νυχτερινής διασκέδασης ξεφυτρώνουν σαν τα μανιτάρια.
- 7 Λείπετε πολύ καιρό από το σπίτι;
- 8 Είναι πολύ ευχάριστος για παρέα.
- 9 Ελπίζω να είσαι συνεπής στο ραντεβού μας.
- 10 Μην κάνεις πολλή φασαρία στο μάθημα.

Liste 9: (Wortverständlichkeit V = 39,8 % ± 6,4 %)

- 1 Me to paramikro sinachaki perni farmaka.
- 2 Tu aresi na kani podilato sto wuno.
- 3 Pies xenes chores echete episkefthi?
- 4 Simmetechi me zilo stis athlitikes ekdilosis tu scholiu tu.
- 5 Ine allergiki stis fraules.
- 6 I efimerides den tha kikloforisun awrio logo argias.
- 7 Pos legete i paralia pu kolibisate?
- 8 O kenurgios diefthintis ine afstiros.
- 9 I agapimeni tu musiki ine i dimotiki.
- 10 Den tha xechaso na su telefoniso.

Κατάλογος 9:

- 1 Με το παραμικρό συναχάκι παίρνει φάρμακα.
- 2 Του αρέσει να κάνει ποδήλατο στο βουνό.
- 3 Ποιές ξένες χώρες έχετε επισκεφθεί;
- 4 Συμμετέχει με ζήλο στις αθλητικές εκδηλώσεις του σχολείου του.
- 5 Είναι αλλεργική στις φράουλες.
- 6 Οι εφημερίδες δεν θα κυκλοφορήσουν αύριο λόγω αργίας.
- 7 Πώς λέγετε η παραλία που κολυμπήσατε;
- 8 Ο καινούργιος διευθυντής είναι αυστηρός.
- 9 Η αγαπημένη του μουσική είναι η δημοτική.
- 10 Δεν θα ξεχάσω να σου τηλεφωνήσω.

Liste 10: (Wortverständlichkeit V = 40 % ± 7,2 %)

- 1 O asthenis ipotropiase ke anewase pireto.
- 2 Dixe sti thia tin zografia pu zografises.
- 3 Pote egine avtos o kawgas?
- 4 O gios mu enthusiazete me tis parastasis tu Karagiozi.
- 5 Kiklofori schedon panta aktenistos.
- 6 I thermo-krasies ine ipsiles gia tin epochi.
- 7 Klisate rantewu me ton giatro sas?
- 8 I musiki afti ine eknewristiki.
- 9 Ta themata tis efimeridas simera ine stenochora.
- 10 Perimeno stin ura gia ta isitiria.

Κατάλογος 10:

- 1 Ο ασθενής υποτροπίασε και ανέβασε πυρετό.
- 2 Δείξε στη θεία την ζωγραφιά που ζωγράφισες.
- 3 Πότε έγινε αυτός ο καυγάς;
- 4 Ο γιος μου ενθουσιάζεται με τις παραστάσεις του Καραγκιόζη.
- 5 Κυκλοφορεί σχεδόν πάντα ακτένιστος.
- 6 Οι θερμο-κρασίες είναι υψηλές για την εποχή.
- 7 Κλείσατε ραντεβού με τον γιατρό σας;
- 8 Η μουσική αυτή είναι εκνευριστική.
- 9 Τα θέματα της εφημερίδας σήμερα είναι στενόχωρα.
- 10 Περιμένω στην ουρά για τα εισιτήρια.

Liste 11: (Wortverständlichkeit V = 40 % ± 7,1 %)

- 1 Stolistike san na pigene se dexiosi.
- 2 Paradechthike me axio-prepia tin itta tu.
- 3 Pote efkerite gia na kuwentiasume?
- 4 To roloi mu pigeni kathe mera ena leptο brosta.
- 5 Pezi me omadiko pneuma.
- 6 Me ti zesti oli xechithikan stis akro-gialies.
- 7 Ti tha eleges gia ena taxidi?
- 8 I porta eklise me thoriwo.
- 9 Tu ekane minisi giati tu milise aschima.
- 10 Xechase tin tsanta tu sto taxi.

Κατάλογος 11:

- 1 Στολίστηκε σαν να πήγαινε σε δεξίωση.
- 2 Παραδέχθηκε με αξιο-πρέπεια την ήττα του.
- 3 Πότε ευκαιρείτε για να κουβεντιάσουμε;
- 4 Το ρολόι μου πηγαίνει κάθε μέρα ένα λεπτό μπροστά.
- 5 Παίζει με ομαδικό πνεύμα.
- 6 Με τη ζέστη όλοι ξεχύθηκαν στις ακρο-γιαλιές.
- 7 Τι θα έλεγες για ένα ταξίδι;
- 8 Η πόρτα έκλεισε με θόρυβο.
- 9 Του έκανε μήνυση γιατί του μίλησε άσχημα.
- 10 Ξέχασε την τσάντα του στο ταξί.

Liste 12: (Wortverständlichkeit V = 40 % ± 6,3 %)

- 1 Tha erthume noritera ap' oti ipologizame.
- 2 Min xechasis na walis nero sta kanarinia.
- 3 Pu akriwos wriskete I trapeza?
- 4 Kathe mina wafi me allo chroma ta mallia tis.
- 5 Tu aresun I katadisis.
- 6 O pappus fora panta kapello otan wgeni exo.
- 7 Ti ora klinun simera ta katastimata?
- 8 Episkeptete taktika ton odont-iatro.
- 9 Enimerose me, se parakalo, otan mathis neotera.
- 10 Fetos to chimona agorasa kenurgio panofori.

Κατάλογος 12:

- 1 Θα έρθουμε νωρίτερα απ' ότι υπολογίζαμε.
- 2 Μην ξεχάσεις να βάλεις νερό στα καναρίνια.
- 3 Πού ακριβώς βρίσκεται η τράπεζα;
- 4 Κάθε μήνα βάφει με άλλο χρώμα τα μαλλιά της.
- 5 Του αρέσουν οι καταδύσεις.
- 6 Ο παππούς φορά πάντα καπέλο όταν βγαίνει έξω.
- 7 Τι ώρα κλείνουν σήμερα τα καταστήματα;
- 8 Επισκέπτεται τακτικά τον οδοντ-ίατρο.
- 9 Ενημέρωσέ με, σε παρακαλώ, όταν μάθεις νεώτερα.
- 10 Φέτος το χειμώνα αγόρασα καινούργιο πανωφόρι.

Liste 13: (Wortverständlichkeit V = 40 % ± 6,7 %)

- 1 Den tu aresun ta kaftera fagita.
- 2 Ithelan poli na nikisun ke ta kataferan.
- 3 Pos legete to psilotero wuno?
- 4 I kinisi stus dromus afxanete kathimerina olo ke perissotero.
- 5 Tha su telefoniso argotera.
- 6 O anipsios mu epithimi na taxidefsi sto diastima.
- 7 Pu stekosaste otan egine to distichima?
- 8 Kani monos tu to giatro.
- 9 Afto to aftokinito kinite me iliaki energia.
- 10 Ine protoporos stin prospathia tis anadasosis.

Κατάλογος 13:

- 1 Δεν του αρέσουν τα καυτερά φαγητά.
- 2 Ήθελαν πολύ να νικήσουν και τα κατάφεραν.
- 3 Πώς λέγεται το ψηλότερο βουνό;
- 4 Η κίνηση στους δρόμους αυξάνεται καθημερινά όλο και περισσότερο.
- 5 Θα σου τηλεφωνήσω αργότερα.
- 6 Ο ανηψιός μου επιθυμεί να ταξιδεύσει στο διάστημα.
- 7 Πού στεκόσαστε όταν έγινε το δυστύχημα;
- 8 Κάνει μόνος του το γιατρό.
- 9 Αυτό το αυτοκίνητο κινείται με ηλιακή ενέργεια.
- 10 Είναι πρωτοπόρος στην προσπάθεια της αναδάσωσης.

Liste 14: (Wortverständlichkeit V = 40 % ± 6,6 %)

- 1 To tzami ragise apo to chtipima.
- 2 O pappus pigeni kathe proi sto kafenio.
- 3 Ti ora archizi I parastasi?
- 4 Ascholite me tis ores me tin kathariotita tu spitiu.
- 5 Protima tis xines gefsis.
- 6 Psonizi apo tin laiki agora schedon kathe ewdomada.
- 7 Boris na mu danisis mia obrella ?
- 8 To kalokeri i meres megalonun.
- 9 O mathitis diabazi poli gia tis exetasis.
- 10 To Sawato-kiriako kimate ligo perissotero.

Κατάλογος 14:

- 1 Το τζάμι ράγισε από το χτύπημα.
- 2 Ο παππούς πηγαίνει κάθε πρωί στο καφενείο.
- 3 Τι ώρα αρχίζει η παράσταση;
- 4 Ασχολείται με τις ώρες με την καθαριότητα του σπιτιού.
- 5 Προτιμά τις ξινές γεύσεις.
- 6 Ψωνίζει από την λαϊκή αγορά σχεδόν κάθε εβδομάδα.
- 7 Μπορείς να μου δανείσεις μια ομπρέλλα;
- 8 Το καλοκαίρι οι μέρες μεγαλώνουν.
- 9 Ο μαθητής διαβάζει πολύ για τις εξετάσεις.
- 10 Το Σαββατο-κύριακο κοιμάται λίγο περισσότερο.

Liste 15: (Wortverständlichkeit V = 40 % ± 6,8 %)

- 1 I odigisi sta chionia ine epikindini.
- 2 Psachnun na wrun ena spiti me isichia.
- 3 Apo pote echete na sinantithite?
- 4 Anarotieme an tha borusame na ftiaxume ena thermo-kipio.
- 5 Pezi klarino sta panigiria.
- 6 Den echi diathesi na giortasi ta genethlia tu.
- 7 Giati den thelis na me akusis?
- 8 To fthinoporo ta chelidonia apodimun.
- 9 Kalitera na minete edo, epidi ine arga.
- 10 Akriwos edo borume na agorasume paputsia.

Κατάλογος 15:

- 1 Η οδήγηση στα χιόνια είναι επικίνδυνη.
- 2 Ψάχνουν να βρουν ένα σπίτι με ησυχία.
- 3 Από πότε έχετε να συναντηθείτε;
- 4 Αναρωτιέμαι αν θα μπορούσαμε να φτιάξουμε ένα θερμο-κήπιο.
- 5 Παίζει κλαρίνο στα πανηγύρια.
- 6 Δεν έχει διάθεση να γιορτάσει τα γενέθλιά του.
- 7 Γιατί δεν θέλεις να με ακούσεις;
- 8 Το φθινόπωρο τα χελιδόνια αποδημούν.
- 9 Καλύτερα να μείνετε εδώ, επειδή είναι αργά.
- 10 Ακριβώς εδώ μπορούμε να αγοράσουμε παπούτσια.

Liste 16: (Wortverständlichkeit V = 40 % ± 6,4 %)

- 1 Pare mazi su ke to magio.
- 2 I karekla espase apo to waros tu.
- 3 Pos perasate to Sawato-kiriako?
- 4 Tu aresi na wlepi ston kinimato-grafo erga epochis.
- 5 Efcharisto gia to doro.
- 6 Gia na taxidepsume stin ethniki odo plironume diodia.
- 7 Poso kostizi to mathima ton agglikon?
- 8 Pathiazete poli me tin politiki.
- 9 Ine poli charumenos giati petiche stis exetasis.
- 10 Kathe proi pini ena potiri portokalada.

Κατάλογος 16:

- 1 Πάρε μαζί σου και το μαγικό.
- 2 Η καρέκλα έσπασε από το βάρος του.
- 3 Πώς περάσατε το Σαββατο-κύριακο;
- 4 Του αρέσει να βλέπει στον κινηματο-γράφο έργα εποχής.
- 5 Ευχαριστώ για το δώρο.
- 6 Για να ταξιδέψουμε στην εθνική οδό πληρώνουμε διόδια.
- 7 Πόσο κοστίζει το μάθημα των αγγλικών;
- 8 Παθιάζεται πολύ με την πολιτική.
- 9 Είναι πολύ χαρούμενος γιατί πέτυχε στις εξετάσεις.
- 10 Κάθε πρωί πίνει ένα ποτήρι πορτοκαλάδα.

Liste 17: (Wortverständlichkeit V = 40 % ± 7,6 %)

- 1 Apo afti tin ilikia echi filenada.
- 2 O asti-nomos dierewna mia periergi ipothesi.
- 3 Boris na mu danisis chrimata?
- 4 I afxisi tu tim-arithmu prowlimatizi kathe iko-genia.
- 5 Iproferi apo chronio nosima.
- 6 Ine diskolo na grapsi kanis istories gia pedia.
- 7 Posa gatia filixenite stin awli sas?
- 8 To kalokeri chrisimopiume panta antiliako.
- 9 Kathise konta sto tzaki gia na zestathi.
- 10 To sakaki tu echi makria manikia.

Κατάλογος 17:

- 1 Από αυτή την ηλικία έχει φιλενάδα.
- 2 Ο αστυ-νόμος διερευνά μια περίεργη υπόθεση.
- 3 Μπορείς να μου δανείσεις χρήματα;
- 4 Η αύξηση του τιμ-άριθμου προβληματίζει κάθε οικο-γένεια.
- 5 Υποφέρει από χρόνιο νόσημα.
- 6 Είναι δύσκολο να γράψει κανείς ιστορίες για παιδιά.
- 7 Πόσα γιατί φιλοξενείτε στην αυλή σας;
- 8 Το καλοκαίρι χρησιμοποιούμε πάντα αντηλιακό.
- 9 Κάθισε κοντά στο τζάκι για να ζεσταθεί.
- 10 Το σακάκι του έχει μακριά μανίκια.

Liste 18: (Wortverständlichkeit V = 39.8 % ± 6,2 %)

- 1 Den xero tipote gia to thema.
- 2 To taxi stamatisse paranoma ke prokalese distichima.
- 3 Pote archizun xana ta mathimata?
- 4 Tha boruses na ton woithisis na kuwalisi tis aposkewes.
- 5 O uranos ine xasteros.
- 6 Tin alli ebdomada tha kopso ta mallia mu.
- 7 Erchese mazi mu ston kinimato-grafo?
- 8 I afirmada tu ine parimiodis.
- 9 Megalonontas theli na ascholithi me tis kataskewes.
- 10 Afti tin epochi apagorewete to kinigi.

Κατάλογος 18:

- 1 Δεν ξέρω τίποτε για το θέμα.
- 2 Το ταξί σταμάτησε παράνομα και προκάλεσε δυστύχημα.
- 3 Πότε αρχίζουν ξανά τα μαθήματα;
- 4 Θα μπορούσες να τον βοηθήσεις να κουβαλήσει τις αποσκευές.
- 5 Ο ουρανός είναι ξάστερος.
- 6 Την άλλη εβδομάδα θα κόψω τα μαλλιά μου.
- 7 Έρχεσαι μαζί μου στον κινηματο-γράφο;
- 8 Η αφηρημάδα του είναι παροιμιώδης.
- 9 Μεγαλώνοντας θέλει να ασχοληθεί με τις κατασκευές.
- 10 Αυτή την εποχή απαγορεύεται το κυνήγι.

Liste 19: (Wortverständlichkeit V = 40 % ± 7,1 %)

- 1 Katharizi to rodi me megali epimelia.
- 2 Siga - siga I diarkia ton epochon allazi.
- 3 Den nomizete oti kani krio?
- 4 Ta mnimia tis archeotitas thawmazonte gia tin teliotita tus.
- 5 Pistewi oti wlepi fantasmata.
- 6 Echete ipomoni, o serwitoros tha erthi se ligo.
- 7 Giati den akus ton patera su?
- 8 Mu aresun poli ta amigdala.
- 9 Prin kimithi pini ena potiri zesto gala.
- 10 To skafos aniki ston kubaros mu.

Κατάλογος 19:

- 1 Καθαρίζει το ρόδι με μεγάλη επιμέλεια.
- 2 Σιγά - σιγά η διάρκεια των εποχών αλλάζει.
- 3 Δεν νομίζετε ότι κάνει κρύο;
- 4 Τα μνημεία της αρχαιότητας θαυμάζονται για την τελειότητά τους.
- 5 Πιστεύει ότι βλέπει φαντάσματα.
- 6 Έχετε υπομονή, ο σερβιτόρος θα έρθει σε λίγο.
- 7 Γιατί δεν ακούς τον πατέρα σου;
- 8 Μου αρέσουν πολύ τα αμύγδαλα.
- 9 Πριν κοιμηθεί πίνει ένα ποτήρι ζεστό γάλα.
- 10 Το σκάφος ανήκει στον κουμπάρο μου.

Liste 20: (Wortverständlichkeit V = 40 % ± 7,1 %)

- 1 I zewres miazun me ta aloga.
- 2 Ta gialia su ine pano sto trapezi.
- 3 Pios efage to fagito mu?
- 4 Tora sta geramata endiaferete gia tin istero-fimia tu.
- 5 Charikame poli pu irthate.
- 6 O tsopanis ewgale ta prowata apo to mantri.
- 7 Xeris pu meni o filis mu?
- 8 Xafnika emfanistike to uranio toxo.
- 9 To gramma su mu edose megali chara.
- 10 Prospathi na skarfalosi stin psiloteri korifi.

Κατάλογος 20:

- 1 Οι ζέβρες μοιάζουν με τα άλογα.
- 2 Τα γυαλιά σου είναι πάνω στο τραπέζι.
- 3 Ποιός έφαγε το φαγητό μου;
- 4 Τώρα στα γεράματα ενδιαφέρεται για την υστερο-φημία του.
- 5 Χαρήκαμε πολύ που ήρθατε.
- 6 Ο τσοπάνης έβγαλε τα πρόβατα από το μαντρί.
- 7 Ξέρεις πού μένει ο φίλος μου;
- 8 Ξαφνικά εμφανίστηκε το ουράνιο τόξο.
- 9 Το γράμμα σου μου έδωσε μεγάλη χαρά.
- 10 Προσπαθεί να σκαρφαλώσει στην ψηλότερη κορυφή.

Danksagung

Herrn Professor Dr. med. Jan Helms, dem langjährigen Direktor der HNO-Klinik der Universität, bin ich zu großem Dank verpflichtet. Er hat den Gastaufenthalt an der Würzburger Klinik ermöglicht, mir freundlicherweise das Thema überlassen und in großzügiger Weise die Möglichkeit geboten, die vorzügliche Ausstattung der Klinik freizügig zu nutzen. Ohne diese Unterstützung wäre diese Arbeit nicht möglich gewesen.

Bei Herrn Professor Dr. med. Rudolf Hagen, Direktor der Klinik, und Herrn Priv.-Doz. Dr. Ing. habil. Mario Cebulla möchte ich mich für die freundliche Übernahme der Referate bedanken.

Herrn Prof. Dipl.-Ing. Moser gilt mein Dank für die hervorragende Betreuung bei dieser Arbeit, sein förderndes Interesse und für die viele Geduld bei der Beantwortung meiner Fragen und die Assistenz bei der Auswertung.

Herr Dr. Franz Schön, als Physiker mit der Betreuung der CI-Patienten betraut, hat in dieser Funktion die CI-Patienten ausgiebig mit dem HSM-Test getestet. Ihm danke ich für viele Tipps im Umgang mit Microsoft Excel und Word, seine Unterstützung bei der statistischen Auswertung, seine programmtechnische Hilfe bei der Zusammenstellung der endgültigen Listen und nicht zuletzt für seine korrigierenden Eingriffe, wenn sich die deutsche Sprache für die Griechin als schwierig zu handhabende Materie erwies.

Herrn Dipl.-Ing. Stefan Brill danke ich für seine hilfreichen Anregungen bei der Planung der Versuche.

Herrn Dipl.-Ing. Schmitt von der Firma MED-EL schulde ich Dank für die digitale Bearbeitung des Sprachmaterials und die Generierung des Rauschsignals.

Herr Dipl.-Ing. Alexander Möltner hat dafür gesorgt, dass die Technik in der großen Audiometrikabine der Klinik einwandfrei funktionierte, und wenn nicht, dann konnte er Abhilfe schaffen. Er hat das Computerprogramm zur Durchführung der Sprachverständlichkeitstests für mich adaptiert. Für all dies und seine stete freundliche Aufnahme bin ich ihm zu großem Dank verpflichtet.

Darüber hinaus danke allen Mitarbeitern der Klinik, die ganz unscheinbar durch eine kleine Hilfestellung das eine oder andere Hindernis aus dem Weg räumten.

Obwohl ich sie nicht namentlich nennen kann, bewirkten sie, dass ich mich an den Aufenthalt in der Klinik in angenehmer Weise erinnere.

Nicht zuletzt danke ich allen Probanden für die vielen mir und der Wissenschaft geopfert Stunden, ihre Ausdauer und Bereitschaft konzentriert bei der Sache zu bleiben, auch wenn es schwerfiel.