

# Einfluss des Störgeräusches auf das Sprachverstehen von Einsilbern

## *Effect of noise on speech recognition of monosyllables*

Alexandra Winkler und Inga Holube

Institut für Hörtechnik und Audiologie, Jade Hochschule und Exzellenzcluster „Hearing4All“, Oldenburg

**Zusammenfassung:** In der Hilfsmittelrichtlinie sind die Indikationskriterien für eine Hörgeräteversorgung und deren Verifikation festgelegt. Diese Verifikation sollte im Störschall stattfinden, da insbesondere störgeräuschbehaftete Hörsituationen für Menschen mit Hörbeeinträchtigung herausfordernd sind. Als mögliches Verfahren bietet sich nach Hilfsmittelrichtlinie u.a. der Freiburger Einsilbertest an. Welches Störgeräusch und welche Lautsprecherdarbietung verwendet werden sollen, ist jedoch in der Hilfsmittelrichtlinie nicht festgelegt. Diese Studie untersuchte, inwiefern das Sprachverstehen von Probanden mit normalem Hörvermögen vom verwendeten Störgeräusch abhängt. Neben dem Freiburger Einsilbertest wurde als Vergleich der Einsilberreimtest nach von Wallenberg und Kollmeier verwendet. Als Störgeräusche wurden jeweils ein aus dem Sprachmaterial generiertes Rauschen, das CCITT-Rauschen, das ICRA1-Rauschen und das Verdeckungsgeräusch nach DIN 60645-2 bei verschiedenen Signal-Rausch-Abständen aus der Vorne-Richtung dargeboten. Für jeden Sprachtest wurden störgeräuschabhängige Diskriminationsfunktionen angepasst und der Signal-Rausch-Abstand für ein Sprachverstehen von 50 % ( $L_{50}$ ) mit dazugehöriger Steigung bestimmt. Die störgeräuschabhängigen  $L_{50}$ -Werte unterschieden sich für jeden Sprachtest signifikant und wiesen einen Unterschied von ca. 6 dB zwischen den beiden Sprachtests auf. Die Verdeckungswirkung der verschiedenen Störgeräusche konnte nur teilweise mit dem Speech Intelligibility Index abgebildet werden. Im Vergleich zu Literaturwerten für die entsprechenden Sprachtests in Ruhe wurde eine höhere Steigung beobachtet. Aufgrund der signifikant unterschiedlichen  $L_{50}$ -Werte wird empfohlen, die Hilfsmittelrichtlinie genauer zu formulieren, mindestens jedoch die Angabe des Störgeräusches zu fordern.

**Stichwörter:** Sprachaudiometrie, Freiburger Einsilbertest, Einsilberreimtest, Störgeräusch, Sprachverständlichkeitsschwelle, Speech Intelligibility Index

**Abstract:** The guidelines for hearing aid treatment include criteria for hearing aid evaluation. Since noisy situations are often very challenging for people with hearing impairment, speech recognition tests in noise are essential. Hearing aid evaluation in noise can be carried out using the Freiburg monosyllabic speech test, but information on which noise is to be applied is missing in the guidelines. Therefore, this study focused on the impact of different noise signals on monosyllabic speech recognition. The Freiburg monosyllabic speech test and the monosyllabic rhyme test of von Wallenberg and Kollmeier were used as speech tests. Speech-shaped noise that had been developed out of the speech material, CCITT, ICRA1 and a noise according to DIN 60645-2 were chosen as noise signals. Both, noise and speech signals, were presented from one loudspeaker in front of the listener at different signal-to-noise-ratios. Noise dependent discrimination functions were fitted and speech recognition thresholds for 50 % correct were estimated for both speech tests. Statistical analysis showed significant difference in speech recognition among the noises. The thresholds for the two speech tests differed by approx. 6 dB. Modelling the measurement outcomes with the Speech Intelligibility Index only partially reflected the influence of the different noises on speech recognition. Compared to literature data for speech tests in quiet, steeper discrimination functions were obtained independently of the different speech tests and noise. Due to significant differences in speech recognition thresholds, the guidelines for hearing-aid treatment should declare more precisely the test conditions or should at least demand to specify the used noise type.

**Keywords:** Speech audiometry, Freiburg monosyllabic speech test, monosyllabic rhyme test, noise, speech recognition threshold, Speech Intelligibility Index

**Korrespondenzadresse:**

Alexandra Winkler  
 Institut für Hörtechnik und Audiologie  
 Jade Hochschule  
 Ofener Str. 16/19  
 26121 Oldenburg  
 alexandra.winkler@jade-hs.de

## Einleitung

Die zwischenmenschliche Kommunikation ist ein wesentlicher Bestandteil des täglichen Lebens und kann für Menschen mit Hörverlust beeinträchtigt sein. Situationen mit Hintergrundgeräuschen sind eine besonders große Herausforderung für diese Menschen, da sie oft Probleme haben, die Sprache im Störgeräusch zu verstehen (Cocktailparty-Effekt). Diese Beeinträchtigungen sollen durch eine Hörgeräteversorgung bestmöglich ausgeglichen werden.

Wie der Erfolg einer Hörgeräteversorgung im Störschall bemessen werden kann, ist unter anderem in der Hilfsmittelrichtlinie (2017) geregelt. Nach der jetzigen Fassung der Hilfsmittelrichtlinie vom 17.02.2017 kann u.a. der Freiburger Einsilbertest (FBE) nach Hahlbrock (1953) für die Erfolgskontrolle der Hörgeräteversorgung im Störschall verwendet werden. Die Verbesserung im Sprachverstehen muss danach mindestens 10 %-Punkte im Vergleich zur unversorgten Messung betragen. Angaben zur räumlichen Konfiguration der Lautsprecher und eine Beschreibung des zu verwendenden Störgeräusches fehlen jedoch in der Hilfsmittelrichtlinie.

In dieser Studie wurde der FBE bei Verwendung von verschiedenen Störgeräuschen durchgeführt und mit einem für Störgeräuschmessungen optimierten Sprachtest, dem Reimtest nach von Wallenberg und Kollmeier (WAKO, 1989), verglichen. Der WAKO verwendet ebenfalls Einsilber, die jedoch nicht wie beim FBE auf einsilbige Substantive beschränkt sind. Der WAKO ist im Gegensatz zum FBE ein geschlossener Test. Der Proband bekommt fünf Antwortalternativen visuell präsentiert, aus denen er das verstandene Wort auswählt. Müller (1992) evaluierte den WAKO mit Normalhörenden und Holube (1993) nutzte den WAKO, um den Einfluss verschiedener räumlicher Richtungen von Störgeräusch und Sprache auf das Sprachverstehen zu untersuchen. Holube (1993) erreichte bei einem Signal-Rausch-Abstand (SNR) von -11,9 dB ein Sprachverstehen von 50 % ( $L_{50}$ ) bei frontaler Darbietung. Dass der WAKO als Alternative für den FBE in Ruhe eingesetzt werden kann, zeigten bereits Sukowski et al. (2009, 2010) im Zuge der Untersuchung von Begutachtungsverfahren. Dabei ergab sich für gleiches Sprachverstehen eine Pegelreduktion beim WAKO im Vergleich zum FBE um 15 dB. Der Unterschied zwischen beiden Sprachtests im Störgeräusch ist bisher nicht bekannt.

Der FBE wurde in der Vergangenheit unter verschiedenen Konditionen (z. B. Darbietungsart, Hörvermögen der Probanden) im Störschall untersucht (Winkler und Holube 2014). Keine dieser Studien erfüllte die Anforderungen der DIN EN ISO 8253-3 (2012). Diese Norm befasst sich unter anderem mit der Evaluierung und Validierung von Sprachtests im Störgeräusch. Die Norm empfiehlt die Darbietung eines aus dem Sprachmaterial generierten Rauschens, aber auch andere frequenzbewertete, nicht modulierte Störgeräusche werden genannt. Diese sind das Verdeckungsgeräusch nach DIN 60645-2 (1997), das CCITT-Rauschen nach ITU-T Recommendation G.227 (1993) und das ICRA1-Rauschen (Dreschler et al. 2001). Nach Festen und Plomp (1990) besteht die beste Verdeckungswirkung, wenn Störgeräusch und Sprachmaterial das gleiche mittlere Langzeitspektrum aufweisen. Das zum WAKO gehörige Rauschen wurde deshalb durch Aufaddieren der einsilbigen Wörter erstellt (Müller 1992). Durch unterschiedlich lange Pausen zwischen den Wörtern und Anfangsoffsets wurde eine Synchronisation und Periodizität zwischen den Wörtern weitgehend minimiert. Aus dem Sprachmaterial des FBE wurde von Döring und Hamacher (1992) das Dreinsilberrauschen (auch als Döringrauschen bekannt) generiert. Die Wörter des FBE wurden ohne Pause

32-fach zeitversetzt überlagert. Zusätzlich wurde dem Dreinsilberrauschen Hall beigefügt, um den Cocktailparty-Effekt und eine natürlichere Störgeräuschsituation zu simulieren. Holube et al. (2017) analysierten verschiedene Störgeräusche in Relation zum Sprachmaterial des FBE. Daraus ergab sich die beste spektrale Übereinstimmung, wie erwartet, mit dem Dreinsilberrauschen, aber auch das ICRA1-Rauschen zeigte eine gute spektrale Übereinstimmung.

Als ein objektives Maß für das Sprachverstehen im Störgeräusch kann der Speech Intelligibility Index (SII, ANSI S3.5, 1997) verwendet werden. Verschiedene Untersuchungen (z. B. Festen und Plomp 1990, Holube und Kollmeier 1996, Beutelmann und Brand 2006, Rhebergen et al. 2010, Lazarus et al. 2007) nutzten den SII oder seinen Vorgänger, den Articulation Index, um das Sprachverstehen in verschiedenen Hörsituationen vorherzusagen. Diese Modell-Maße verdeutlichen den Einfluss des SNR auf das Sprachverstehen. Der SII berechnet unter Verwendung der Langzeitspektren von Sprache und Störgeräusch die Hörbarkeit in jedem Frequenzband in Abhängigkeit von Hörvermögen und Darbietungspegel. Da nicht der gesamte Frequenzbereich für das Sprachverstehen die gleiche Relevanz besitzt, erfolgt eine Gewichtung der einzelnen Frequenzbänder in Abhängigkeit vom verwendeten Sprachmaterial. Je nach SNR wird der Beitrag jedes einzelnen Bandes zur Hörbarkeit berechnet und zu einem Wert zwischen 0 und 1 gewichtet aufsummiert, wobei 1 auf ein sehr gutes und 0 auf kein Sprachverstehen hindeutet. Nach Magnusson (1996) liegt der SII für 50 %iges Einsilberverstehen zwischen 0,2 und 0,3. Dieser Wertebereich dient jedoch nur als grobe Orientierung, da er vom verwendeten Sprachmaterial abhängt. Mit dem SII kann untersucht werden, ob die unterschiedliche Verdeckungswirkung von Störgeräuschen, die perzeptive Unterschiede im Sprachverstehen bewirkt, objektivierbar ist.

Dieser Beitrag zeigt zum einen den Einfluss des verwendeten Störgeräusches auf das Sprachverstehen beim FBE und beim WAKO und analysiert, inwiefern die Unterschiede zwischen den Störgeräuschen mittels SII modelliert werden können. Zum anderen wird das Sprachverstehen beider Sprachtestverfahren im Störgeräusch miteinander und mit Literaturdaten verglichen. Diese Untersuchung soll dazu beitragen, die Vorgaben zur Durchführung des FBE im Störgeräusch datenbasiert zu konkretisieren.

## Methoden

### Probanden

An den Messungen nahmen 40 Probanden (24 w, 16 m) im Alter von 18 bis 25 Jahren (Median: 22 Jahre) teil. Das Hörvermögen wurde seitengetreten ermittelt. Dabei war für jedes Ohr ein Hörvermögen von maximal 10 dB HL im Frequenzbereich von 250 Hz bis 8 kHz zulässig, wobei bei höchstens zwei Frequenzen ein Hörverlust von maximal 15 dB HL erlaubt war. Das Alter und das Hörvermögen entsprachen den Vorgaben der DIN EN ISO 8253-3 (2012). Die tonaudiometrischen Hörschwellen wurden mit dem Audiometer Siemens Unity 2 und dem Kopfhörer Sennheiser HDA 200 nach DIN 8253-1 (2011) ermittelt. Die Anforderung an die otologische Normalität (Gesundheitszustand, Lärmeinwirkung, Einnahme ototoxischer Wirkstoffe, erblich bedingter Hörverlust, Ohrkrankheiten) nach DIN EN ISO 8253-1 (2011) wurde durch eine mündliche Befragung der Probanden überprüft. Allen Probanden war das verwendete Sprachmaterial unbekannt. Für die Teilnahme an der Studie erhielten alle Probanden eine Aufwandsentschädigung in Höhe von 10 € pro Stunde. Die Studie wurde von der Kommission für Forschungs-

folgenabschätzung und Ethik der Carl von Ossietzky Universität Oldenburg (Drs. 1/2015) genehmigt.

### Freiburger Einsilbertest

Der FBE besteht aus 20 Testlisten mit jeweils 20 einsilbigen Nomen (Hahlbrock 1953). Für die Untersuchung wurde die Aufnahme von 1969 (Brinkmann 1974) von der Siemens-CD mit der Sach-Nr. 7970155 HH 922 verwendet. Die Reihenfolge der Wortpräsentationen innerhalb der Testlisten entsprach DIN 45621-1 (1995). Da die Listen 5, 11, 12 und 15 nach Baljic et al. (2016) in Ruhe nicht perzeptiv äquivalent waren, wurden diese Listen nicht verwendet. Das prozentuale Sprachverstehen  $p$  ergab sich aus der Anzahl der richtig verstandenen Wörter pro Liste in Relation zur Gesamtanzahl (20 Wörter) pro Liste. Während der Messung saß der Versuchsleiter seitlich hinter dem Probanden, um die verstandenen Wörter über einen Touchscreen einzugeben.

### Einsilberreimtest

Für den WAKO wurde die gekürzte Version (25 Wörter pro Testliste) nach Brand und Wagener (2005) genutzt, der 18 Testlisten mit 321 unterschiedlichen Wörtern beinhaltet. Die Wörter unterschieden sich im Anlaut, Inlaut oder Auslaut. Vor jedem Wort wurde der Satz: „Bitte markieren Sie das Wort“ gesprochen. Auf einem Touchscreen mussten die Probanden das verstandene Wort aus fünf Antwortalternativen auswählen. Das erreichte prozentuale Sprachverstehen  $p$  wurde um die Ratewahrscheinlichkeit aufgrund der fünf Antwortalternativen korrigiert. Der Versuchsleiter befand sich während der Messung außerhalb der Hörkabine.

### Störgeräusche

Das Sprachverstehen im Störgeräusch wurde mit dem Verdeckungsgeräusch nach DIN 60645-2 (1997), dem ICRA1-Rauschen (Dreschler et al. 2001) und dem CCITT-Rauschen der Siemens-CD (Sach-Nr. 7970155 HH 922) ermittelt. Zusätzlich wurde als Störgeräusch das aus dem jeweiligen Sprachmaterial generierte Rauschen verwendet, für den FBE das Dreinsilberrauschen (DRF; Döring und Hamacher 1992, CD 8 von Westra Elektroakustik GmbH) und für den WAKO das Einsilberrauschen (ERW) nach Müller (1992). So wurde jedes Testverfahren mit vier verschiedenen Störgeräuschen präsentiert. Abbildung 1 zeigt die Differenz der Leistungsdichtespektren zwischen der Sprache und den Störgeräuschen für den WAKO (links) und den FBE (rechts). Alle Signale weisen den gleichen mittleren Langzeitpegel auf. Bei einer positiven Differenz ist das Sprachsignal lauter, bei einer negativen das Störgeräusch.

Die beste Übereinstimmung zwischen Sprache und Störsignal besteht prinzipiell bei den aus dem Wortmaterial generierten Störsignalen (schwarze Kurven). WAKO und ERW weichen dennoch im Hochtonbereich ab 2 kHz (ca. 20 dB bei 8 kHz) voneinander ab. Die Abweichungen zwischen FBE und DRF sind geringer (maximal 5,5 dB im Bereich 125 Hz bis 8 kHz).

Das Störgeräusch startete 0,5 s vor dem Wort beim FBE und 0,5 s vor dem Ankündigungssatz beim WAKO und endete jeweils 0,5 s nach jeder Wortpräsentation. Unabhängig vom Sprachtest erfolgte die Ausgabe eines neuen Wortes erst nach der Eingabe der Antwort des vorherigen Wortes am Touchscreen. Somit war die Pausendauer zwischen den Störgeräuschdarbietungen bei beiden Sprachtests nicht

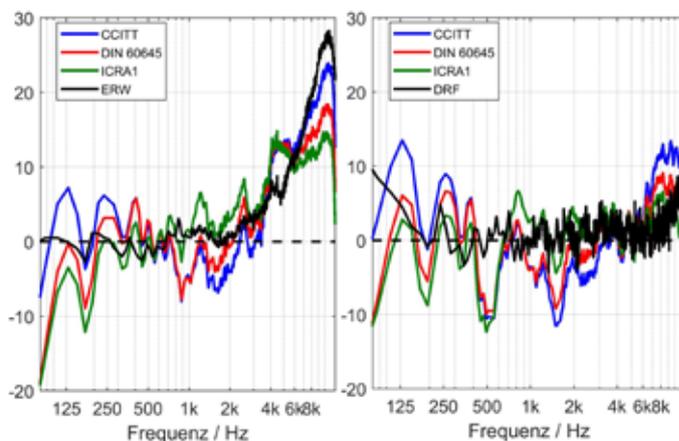


Abbildung 1: Differenz der Leistungsdichtespektren der Sprachsignale und der verschiedenen Störgeräusche für die Einsilber nach von Walzenberg und Kollmeier (WAKO, links) und für die Freiburger Einsilber (FBE, rechts). Alle Signale weisen den gleichen mittleren Langzeitpegel auf.

Figure 1: Difference of power spectral densities of the speech signals and the noises for WAKO (left) and FBE (right). All signals were calibrated to the same average long-term level.

konstant. Der Störgeräuschpegel betrug 60 dB SPL. Der Sprachpegel variierte in Abhängigkeit vom ausgewählten SNR.

### Messverfahren

Eine Diskriminationsfunktion nach Brand und Kollmeier (2002) wurde an das ermittelte prozentuale Sprachverstehen pro SNR angepasst. Gleichung (1) zeigt den mit Hilfe einer Diskriminationsfunktion beschriebenen Zusammenhang zwischen dem prozentualen Sprachverstehen  $p$ , dem dargebotenen SNR ( $L_{SNR}$ ), dem SNR für ein Sprachverstehen von 50 % ( $L_{50}$ ) und der Steigung ( $s_{50}$ ) im dazugehörigen Punkt:

$$p(L_{SNR}, L_{50}, s_{50}) = \frac{100}{1 + e^{-s_{50} \cdot (L_{50} - L_{SNR})}} \quad (1)$$

Um Diskriminationsfunktionen sicher anpassen zu können, sind SNR-Werte erforderlich, die zu einem Sprachverstehen sowohl oberhalb als auch unterhalb von 50 % führen. Angelehnt an die Ergebnisse von Malling (2011) und Holube (1993) wurden folgende SNRs ausgewählt:

- FBE: -6, -3, 0 und 3 dB
- WAKO: -15, -12, -9 und -6 dB.

Jeder Proband hörte jedes der vier Störgeräusche bei jedem der vier ausgewählten SNR-Werte. So wurden ihm für jeden Sprachtest 16 verschiedene Testlisten präsentiert, die sich nicht wiederholten. Die Probanden führten erst alle Messungen für einen Sprachtest durch, bevor zum anderen Sprachmaterial gewechselt wurde. Die eine Hälfte der 40 Probanden begann mit dem FBE und die andere Hälfte mit dem WAKO. Die Probanden hörten zunächst vier Testlisten bei einem SNR und unterschiedlichen Störgeräuschen. Dann wurde

der SNR verändert und die Störgeräusche erneut präsentiert. Dieses Vorgehen wurde für die verbleibenden zwei SNR-Werte übernommen. Die Reihenfolge der SNR-Werte, der Störgeräusche und der Testlisten wurde randomisiert.

## Apparatur

Die Sprachmaterialien und die Störgeräusche beider Testverfahren lagen digital auf der Festplatte des Steuercomputers vor. Die Tonaudiometrie und die Sprachverständlichkeitsmessungen (Oldenburger Messapparatur Forschungsversion 2.0, HörTech gGmbH) fanden in einer gedämmten Freifeldaudiometrikabine statt. Die Ausgabe der Signale für die Sprachtests erfolgte über ein RME Fireface 400 UC und einen Genelec 8030A Lautsprecher. Der Abstand zwischen Proband und Lautsprecher betrug 1,1 m. Die Sprache und das Störgeräusch wurden aus dem Lautsprecher frontal ( $S_0N_0$ ) dargeboten.

## Speech Intelligibility Index

Für die Modellierung des Sprachverstehens wurde der SII nach ANSI S3.5 (1997) verwendet. Dafür erfolgte die Bestimmung der Langzeitpektren der Sprach- und Störsignale in Terzbändern. Für den FBE wurden nur die Einsilber als Sprachsignal verwendet, die auch bei der Messung verwendet wurden. Die Wörter der Listen 5, 11, 12 und 15 wurden also nicht genutzt. Da alle Probanden normalhörend waren, wurde der Hörverlust mit 0 dB HL für das gesamte Frequenzspektrum (160 Hz bis 8 kHz) angenommen. Für die Modellierung wurde das Störgeräusch digital so kalibriert, dass es einem Darbietungspegel von 60 dB SPL entsprach. Der Schalldruckpegel des Sprachmaterials variierte je nach SNR. In der Regel beträgt bei Sprachtests der gemittelte Langzeitpegel des Kalibriersignals demjenigen des Sprachmaterials. Dies trifft für den WAKO und das ERW zu. Beim FBE erfolgte allerdings bei Erstellung der Referenzaufnahme die Aussteuerung für CCITT und FBE impulsbewertet, also nach Spitzenpegeln (DIN 45626-1, 1995). Winkler und Holube (2016) zeigten, dass bei der Siemens-CD der Unterschied im mittleren Langzeitpegel zwischen CCITT und Freiburger Einsilbern 6,5 dB betrug. D. h., bei einem SNR von 0 dB und einem gemittelten Langzeitpegel für das CCITT-Rauschen von 60 dB SPL beträgt der mittlere Langzeitpegel der Freiburger Einsilber 53,5 dB SPL. Diese Differenz wurde bei der SNR-abhängigen Berechnung des SII berücksichtigt. Ein weiterer Parameter bei der SII-Berechnung ist die Wahl der Gewichtungsfunktion (ANSI S3.5 1997), die für verschiedene Sprachmaterialien den Beitrag des jeweiligen Terzbandes zum Sprachverstehen beschreibt (siehe Abbildung 2).

Unter der Annahme, dass mit dem SII die Unterschiede im Sprachverstehen bei der Verwendung verschiedener Störgeräusche modelliert werden können, müsste ein monotoner Zusammenhang zwischen dem gemessenen Sprachverstehen und dem SII beobachtbar sein. Da die Gewichtungsfunktionen nicht explizit für den FBE und den WAKO ermittelt wurden und deshalb nicht eindeutig festgelegt sind, erfolgte die Berechnung des SII für alle Gewichtungsfunktionen aus ANSI S3.5 (1997). Für jede Gewichtungsfunktion wurde in erster Näherung eine Regressionsgerade über alle Störgeräusche und SNR-Werte der Form:

$$SII_{Regr} = \text{Steigung} \cdot x + b \quad (2)$$

bestimmt, wobei  $SII_{Regr}$  dem SII anhand der Regressionsgeraden und  $x$  dem Sprachverstehen entsprach. Für jede Gewichtungsfunktion wurde die mittlere quadratische Abweichung zwischen dem vorher-

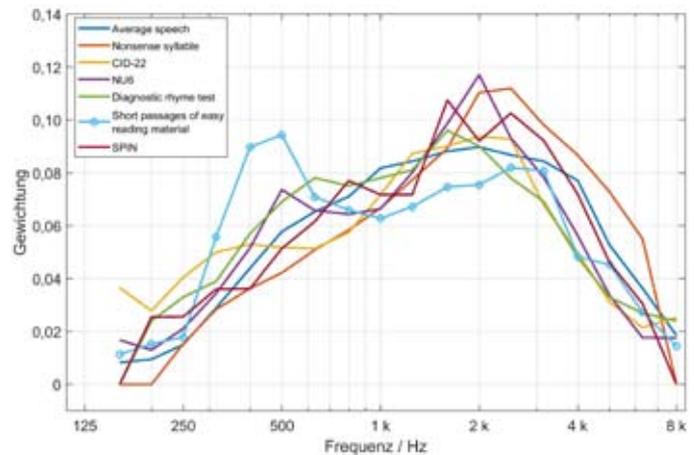


Abbildung 2: Gewichtungsfunktionen für die verschiedenen Sprachmaterialien nach ANSI S3.5 (1997). Die Funktion „Short passages of easy reading material“ ist zur Verbesserung der Sichtbarkeit mit Kreisen dargestellt.

Figure 2: Band importance functions for different speech materials according to ANSI S3.5 (1997). The curve for “Short passages of easy reading material” is marked by circles to improve visibility.

gesagten Wert  $SII_{Modell}$  und dem mit der Regressionsgeraden berechneten  $SII_{Regr}$  bei jeweils gegebenem Sprachverstehen bestimmt. Die Gewichtungsfunktion mit der geringsten mittleren quadratischen Abweichung wurde als beste Anpassfunktion gewählt:

$$\text{Mittelwert} (SII_{Modell} - SII_{Regr})^2 \stackrel{!}{=} \text{minimal} \quad (3)$$

## Statistische Auswertung

Da die probandenspezifischen Werte  $L_{50}$  und  $s_{50}$  nach Shapiro-Wilk-Test nicht normalverteilt waren, erfolgte die statistische Auswertung mit dem Friedman-Test und Post-hoc mit dem Wilcoxon-Test für abhängige Stichproben. Das Signifikanzniveau betrug  $\alpha = 0,05$ . Um  $L_{50}$  und  $s_{50}$  auf signifikante Unterschiede hinsichtlich der verwendeten Störgeräusche zu prüfen, wurde die Anzahl der Paarvergleiche mittels Bonferroni-Korrektur berücksichtigt. Die Boxplots in den Abbildungen 3, 5 und 6 stellen den Median (-), den Interquartilsabstand (Boxlänge), die Whisker (Antennen bis maximal zum 1,5fachen des Interquartilsabstands) und Ausreißer (o) dar. Die gesamte Auswertung erfolgte mit der Software Matlab R2017a der Fa. MathWorks.

## Ergebnisse

Im Folgenden werden zunächst die Ergebnisse im Sprachverstehen für jeden Sprachtest und jedes Störgeräusch dargestellt. Anschließend erfolgen eine statistische Auswertung zur Analyse signifikanter Unterschiede im Sprachverstehen zwischen den verschiedenen Störgeräuschen und die Modellierung mittels SII.

## Sprachverstehen und Diskriminationsfunktionen

In Abbildung 3 ist das Sprachverstehen im WAKO und im FBE in Abhängigkeit vom SNR für die vier verwendeten Störgeräusche darge-

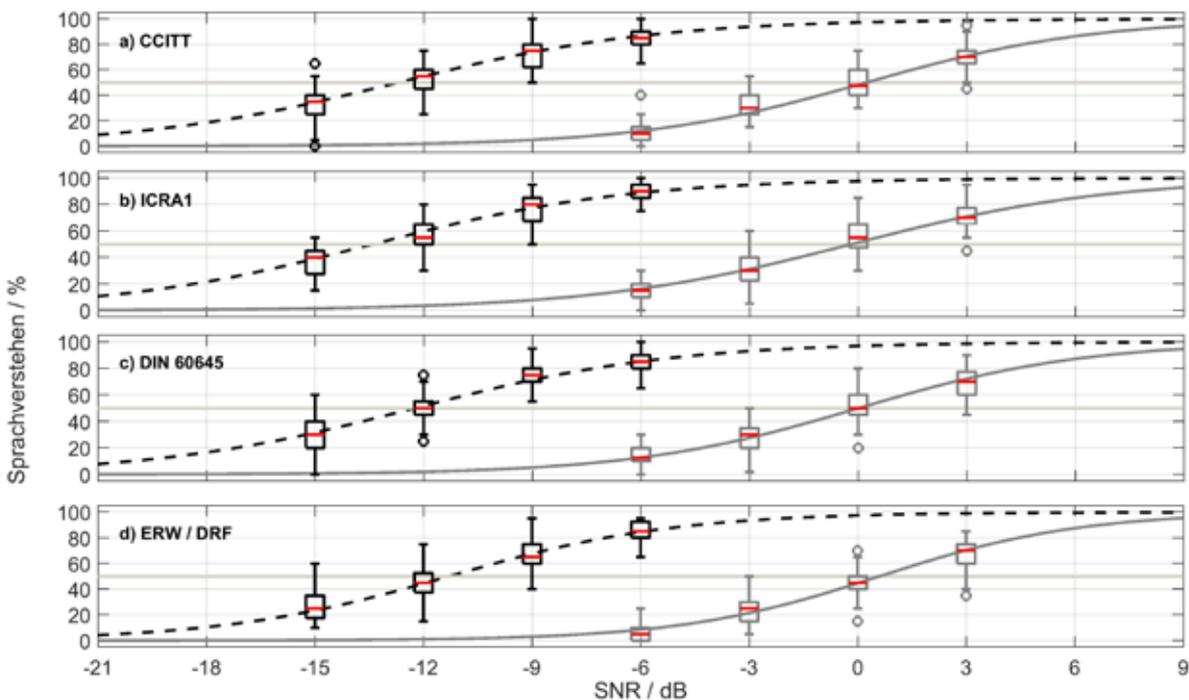


Abbildung 3: Sprachverstehen für WAKO (schwarz) und FBE (grau) in Abhängigkeit vom SNR als Boxplots und die Diskriminationsfunktion anhand der Mediane (rote Linie) im Sprachverstehen pro SNR (WAKO: gestrichelt, FBE: durchgezogen). Für einen einfacheren Vergleich ist bei einem Sprachverstehen von 50 % eine Linie eingezeichnet.

Figure 3: Speech recognition for WAKO (black) and FBE (gray) as boxplots at different SNRs and for different noises. The median speech-recognition scores per SNR (red line) were used for the discrimination functions (WAKO: dashed lines, FBE: solid lines). For a better comparison a line for 50 % speech recognition is drawn.

stellt. Mit Nutzung der Mediane bei jedem SNR wurde zur Illustration für jedes Störgeräusch die Diskriminationsfunktion für den WAKO und den FBE nach Gleichung (1) angepasst.

Aus Abbildung 3 geht hervor, dass die gewählten SNR-Werte den Bereich von 0 bis 100 % Sprachverstehen für die Bestimmung der Diskriminationsfunktionen (mit  $L_{50}$  und  $s_{50}$ ) abdeckten. Unabhängig vom Störgeräusch und vom Sprachtest war die interindividuelle Streuung im Sprachverstehen sehr groß, obwohl alle Probanden normalhörend waren. Der Vergleich zwischen WAKO und FBE bei gleichem Sprachverstehen zeigt einen Unterschied im SNR von ca. 12,5 dB (vgl. Tabelle 1). Bei Berücksichtigung der unterschiedlichen

Kalibrierung der beiden Sprachtests verbleibt ein Unterschied im SNR von ca. 6 dB.

Anhand des Sprachverstehens pro SNR wurden probandenspezifische Diskriminationsfunktionen für jedes Störgeräusch nach Gleichung (1) angepasst und individuelle Werte für  $L_{50}$  und  $s_{50}$  bestimmt. Tabelle 1 enthält die Mediane für die individuellen  $L_{50}$ - und  $s_{50}$ -Werte für jedes Störgeräusch und jeden Sprachtest. Die Diskriminationsfunktionen basierend auf diesen Medianwerten sind zur besseren Vergleichbarkeit der Störgeräusche in Abbildung 4 dargestellt. Die maximale Differenz der störgeräuschspezifischen  $L_{50}$ -Werte war beim WAKO größer (1,8 dB) als beim FBE (0,9 dB).

Tabelle 1: Mediane für  $L_{50}$  und  $s_{50}$  für den FBE und den WAKO in Abhängigkeit vom Störgeräusch. Das Störgeräusch DRF wurde nur beim FBE und das Störgeräusch ERW nur beim WAKO verwendet.

Table 1: Median values of  $L_{50}$  and  $s_{50}$  for FBE and WAKO using different noises. The noise DRF was used for FBE and the noise ERW for WAKO.

Störgeräusch	FBE $L_{50\_Median}$ in dB SNR	FBE $s_{50\_Median}$ in %-Punkte/dB	WAKO $L_{50\_Median}$ in dB SNR	WAKO $s_{50\_Median}$ in %-Punkte/dB
CCITT	0,0	8,0	-12,5	7,0
ICRA1	-0,3	7,5	-13,1	7,0
DIN 60645	0,2	7,5	-12,1	8,0
DRF bzw. ERW	0,6	9,0	-11,3	7,5

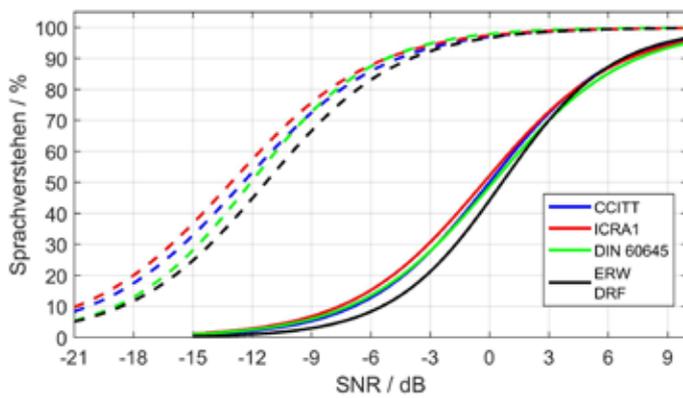


Abbildung 4: Mediane Diskriminationsfunktionen für den WAKO (gestrichelte Linien) und den FBE (durchgezogene Linien) in Abhängigkeit vom verwendeten Störgeräusch.

Figure 4: Median discrimination functions for WAKO (dashed lines) and FBE (solid lines) for the different noises.

**Statistische Unterschiede in  $L_{50}$  und  $s_{50}$**

Abbildung 5 zeigt als Boxplot die Ergebnisse der individuellen  $L_{50}$ - und  $s_{50}$ -Werte für den FBE (obere Reihe) und WAKO (untere Reihe). Ähnlich wie beim Sprachverstehen (vgl. Abbildung 3) war die Streuung für beide Sprachtests und alle Störgeräusche relativ groß. Beim Vergleich der vier Störgeräusche für beide Sprachtests war zunächst zu beobachten, dass der mediane  $L_{50}$  bei dem aus dem Sprachmaterial generierten Rauschen (ERW bzw. DRF) höher lag als bei den anderen Störgeräuschen. Im Folgenden wurden die Unterschiede auf Signifikanz untersucht.

Der Friedman-Test für den FBE ergab signifikante Unterschiede zwischen den störgeräuschabhängigen  $L_{50}$  ( $\chi^2(3) = 21,9, p < 0,001$ ) und  $s_{50}$  ( $\chi^2(3) = 10,18, p = 0,017$ ). Für den FBE sind Unterschiede im  $L_{50}$  zwischen DRF und CCITT, zwischen DRF und ICRA1 sowie zwischen DIN 60645 und ICRA1 signifikant (Post-hoc Wilcoxon-Test, siehe Tabelle 2, orangefarbenes Dreieck). Die Diskriminationsfunktion für DRF wies mit 9,0 %-Punkten pro dB SNR die höchste Steigung auf (siehe Tabelle 1) und unterschied sich signifikant nur zur Steigung von ICRA1 ( $p = 0,040$ ). Alle anderen störgeräuschabhängigen Steigungen waren im Post-hoc-Test nicht signifikant unterschiedlich. Für alle Störgeräusche war die Steigung deutlich höher als für die Ruhemessung (4,6 %-Punkte pro dB, Baljic et al. 2016).

Der Friedman-Test für den WAKO ergab signifikante Unterschiede zwischen den  $L_{50}$  in Abhängigkeit vom Störgeräusch ( $\chi^2(3) = 37,89, p < 0,001$ ). Tabelle 2 zeigt in der Übersicht die p-Werte der Post-hoc-Analyse mit dem Wilcoxon-Test für den  $L_{50}$  im WAKO im grauen Dreieck. Die Unterschiede im  $L_{50}$  zwischen ERW und allen anderen Störgeräuschen sowie zwischen DIN 60645 und ICRA1 sind signifikant. Die Steigung  $s_{50}$  war für die vier Störgeräusche mit 6,9 bis 7,6 %-Punkte pro dB SNR ähnlich und im Friedman-Test nicht signifikant verschieden ( $\chi^2(3) = 4,12, p = 0,249$ ). Im Vergleich zur Steigung in Ruhe (6,6 % Punkte pro dB aus Holube 1993) war die mediane Steigung bei allen Störgeräuschen höher.

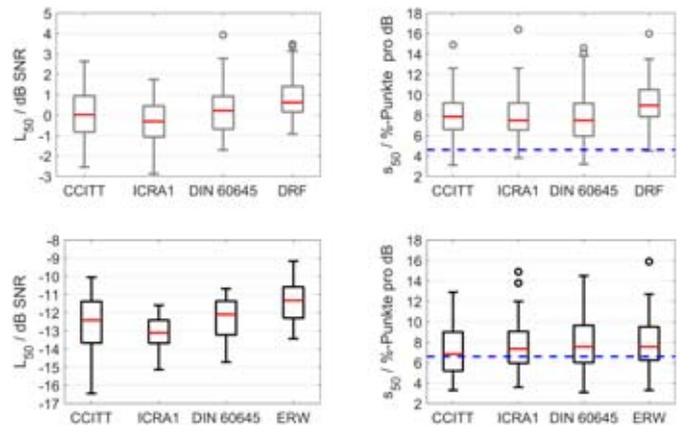


Abbildung 5: Boxplots der individuellen  $L_{50}$  (links) und  $s_{50}$ -Werte (rechts) für den FBE (oben) und den WAKO (unten) für alle Störgeräusche. Die Steigung in Ruhe ist für den FBE ( $s_{50\_Ruhe} = 4,6$  %-Punkte pro dB nach Baljic et al. 2016) und für den WAKO ( $s_{50\_Ruhe} = 6,6$  %-Punkte pro dB nach Holube 1993) als blau gestrichelte Linie dargestellt.

Figure 5: Boxplot of individual  $L_{50}$  (left panel) and  $s_{50}$  (right panel) values for FBE (upper panel) WAKO (lower panel) for all noises. Slope in quiet for FBE ( $s_{50\_quiet} = 4,6$  %-point per dB according to Baljic et al. 2016) and for WAKO ( $s_{50\_quiet} = 6,6$  %-point per dB according to Holube 1993) is shown as a blue dashed line.

Tabelle 2: p-Werte für den Wilcoxon-Test für abhängige Stichproben für den  $L_{50}$  im FBE (oberes Dreieck, orange) und im WAKO (unteres Dreieck, blau). n.s.: nicht signifikant

Table 2: p-values of Wilcoxon test for paired samples for  $L_{50}$ . Top triangle: FBE (orange) bottom triangle: WAKO (blue). n.s.: not significant

	FBE	CCITT	ICRA1	DIN 60645	DRF
WAKO					
CCITT			n.s.	n.s.	0,008
ICRA1		n.s.		0,041	< 0,001
DIN 60645		n.s.	0,002		n.s.
ERW		0,010	< 0,001	0,012	

**Sprachverständlichkeitsvorhersage**

Für jede Gewichtungsfunktion nach ANSI S3.5 (1997) wurde der SII für alle Störgeräusche und beide Sprachsignale bestimmt. In Tabelle 3 sind die mittleren quadratischen Abweichungen der modellierten SII-Werte von der Regressionsgeraden nach Gleichung (3) für den WAKO und den FBE für alle Gewichtungsfunktionen aufgeführt. Alle mittleren quadratischen Abweichungen sind unabhängig von der Gewichtungsfunktion für beide Sprachtests sehr klein. Die geringsten

quadratischen Abweichungen wurden für die Gewichtungsfunktion „Short passages of easy reading material“ („short passages“) erreicht. Bei dieser Gewichtungsfunktion wird im Vergleich zu den anderen Gewichtungsfunktionen der Frequenzbereich von 300 Hz bis 600 Hz

Tabelle 3: Mittlere quadratische Abweichung des SII ( $\cdot 10^4$ ) nach Gleichung 3 in Abhängigkeit von der Gewichtungsfunktion definiert nach ANSI S3.5 (1997).

Table 3: Mean least-squares differences (equation 3) of SII ( $\cdot 10^4$ ) for different band importance functions according to ANSI S3.5 (1997).

	Average speech	Non-sense syllable	CID-22	NU6	Diagnostic rhyme test	Short passages	SPIN
FBE	12,9	16,2	7,2	13,0	9,9	7,0	14,1
WAKO	8,3	10,1	5,3	8,0	6,6	4,4	8,9

stärker und der Frequenzbereich ab 4 kHz geringer gewichtet (siehe Abbildung 2).

Abbildung 6 zeigt die  $SII_{Modell}$ -Werte für die Gewichtungsfunktion „short passages“ in Abhängigkeit vom mittleren Sprachverstehen bei den jeweiligen vier Pegeln für den WAKO (links) und FBE (rechts). Für den WAKO liegen die berechneten  $SII_{Modell}$ -Werte fast ausschließlich auf der Regressionsgeraden und zeigen damit nahezu einen linearen Zusammenhang zum Sprachverstehen. Für den FBE zeigt sich hingegen, dass im Vergleich zum Sprachverstehen die  $SII_{Modell}$ -Werte für DRF zu hoch und für CCITT zu gering sind.

Anhand der Regressionsgeraden kann der SII für ein Sprachverstehen von 50 % ( $SII_{50}$ ) berechnet werden. Der  $SII_{50}$  lag für die Gewichtungs-

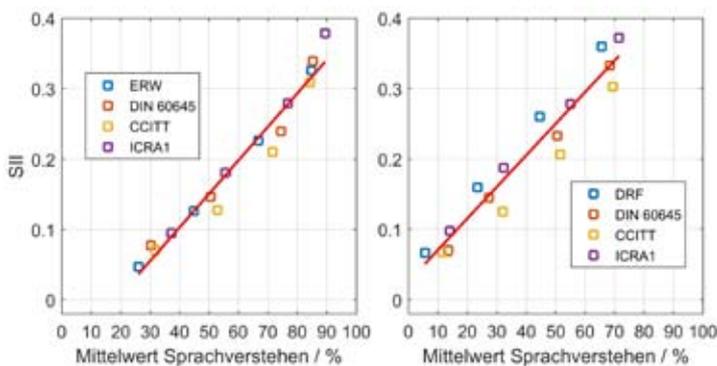


Abbildung 6:  $SII_{Modell}$  in Abhängigkeit vom mittleren Sprachverstehen des WAKO (links) und des FBE (rechts) für die verschiedenen Störgeräusche. Als Gewichtungsfunktion in der SII-Berechnung wurde „Short passages of easy reading material“ nach ANSI S3.5 (1997) verwendet. Die durchgezogene rote Linie gibt die Regressionsgerade an.

Figure 6:  $SII_{Modell}$  as a function of the mean speech recognition in different noises for WAKO (left) and FBE (right). The band importance function „Short passages of easy reading material“ according to ANSI S3.5 (1997) was used for the SII-calculation. The regression line is shown as a red solid line.

funktion „short passages“ für den FBE bei 0,25 und für den WAKO bei 0,15.

## Diskussion

Diese Studie untersuchte den Einfluss des Störgeräusches auf das Einsilberverstehen bei Probanden mit normalem Hörvermögen. Die verwendeten Störgeräusche entsprachen den Empfehlungen der DIN EN ISO 8253-3 (2012) für Sprachtestverfahren. Im Folgenden werden die Ergebnisse der Sprachtests, die spektralen Unterschiede der verwendeten Signale und die Modellierung des Sprachverstehens mittels SII diskutiert.

### Varianz im Sprachverstehen

Die gewählten SNR-Werte resultierten in einem medianen Sprachverstehen unterhalb und oberhalb von 50 % für beide Sprachtests und alle vier Störgeräusche, sodass Diskriminationsfunktionen angepasst werden konnten. Die Varianz im Sprachverstehen bei den SNR-Werten und im  $L_{50}$  war jedoch relativ groß, obwohl alle Probanden ein ähnliches Tonhörvermögen aufwiesen. Diese Abweichungen könnten durch nicht-audiometrische Unterschiede zwischen den Probanden und durch die Messgenauigkeit der Sprachtests bedingt sein. Nicht-audiometrische Gründe sind z.B. die unterschiedliche Bekanntheit der Wörter bei den Probanden und mögliche Schwankungen der Aufmerksamkeit während der Messung.

Die Messgenauigkeit von Sprachtests in Abhängigkeit vom Sprachpegel lässt sich unter Verwendung der verallgemeinerten Binomialverteilung modellieren (Holube et al. 2018). Damit kann die Abweichung des Messwertes vom wahren Wert für das Sprachverstehen abgeschätzt werden. Diese Abweichung beträgt bei Verwendung von Testlisten mit 20 Wörtern beim FBE bis zu  $\pm 15$  %-Punkten (95%-Konfidenzintervall bei einem Sprachverstehen von 50 %). Bei der Modellierung mit der verallgemeinerten Binomialverteilung wird jedoch vorausgesetzt, dass alle Testlisten gleich verständlich sind. Beim FBE wurden vier Testlisten, die beim Sprachverstehen in Ruhe auffällig waren (Baljic et al. 2016), nicht verwendet. Ob jedoch im Störgeräusch die gleichen Testlisten besser oder schlechter verstanden werden, ist bisher noch ungeklärt. Eine mögliche Ungleichheit der Testlisten könnte also zusätzlich zur Varianz beitragen. Um die Messgenauigkeit zu erhöhen, empfiehlt die Hilfsmittelrichtlinie die Verwendung von zwei Testlisten des FBE. Damit reduziert sich das Konfidenzintervall um den Faktor Wurzel 2. Der WAKO besitzt in der verkürzten Version eine ähnliche Anzahl an Wörtern pro Testliste, so dass ein vergleichbares Konfidenzintervall angenommen werden kann. Um die Messgenauigkeit für den WAKO zu verbessern, stehen Testlisten mit 47 oder 72 Wörtern zur Verfügung.

Das Sprachverstehen im Störgeräusch wird typischerweise durch den Schwellenwert  $L_{50}$  charakterisiert. Um den  $L_{50}$ -Wert möglichst genau bestimmen zu können, muss die Diskriminationsfunktion möglichst steil sein (Wagener et al. 1999, Wesselkamp et al. 1992). So bewirken kleine Änderungen des SNR eine möglichst große Veränderung des Sprachverstehens. Im Vergleich zum Sprachverstehen in Ruhe (Baljic et al. 2016, Holube, 1993) wurde für alle Messkonditionen im Störgeräusch eine steilere Diskriminationsfunktion erreicht. Dieser Unterschied zwischen dem Sprachverstehen in Ruhe und im Störgeräusch wurde bereits von Wagener (2004) beschrieben. Wagener (2004) führte dies auf die spektrale Verdeckungswirkung der Störgeräusche

und auf das Verstehen der einzelnen Wörter einer Testliste zurück. Je ähnlicher das Wortverstehen innerhalb einer Testliste ist, desto steiler ist die Diskriminationsfunktion. Für den FBE zeigten Holube et al. (2018), dass das Verstehen der Wörter in Ruhe stark streut. Aufgrund der höheren Steigung kann also angenommen werden, dass die Verständlichkeit der Wörter des FBE im Störgeräusch homogener ist als in Ruhe. Eine Analyse des Wortverstehens im Störgeräusch steht jedoch noch aus.

### Freiburger Einsilbertest

Beim Vergleich der Leistungsdichtespektren der Störgeräusche zum Sprachmaterial fällt auf, dass das DRF, obwohl aus dem Sprachmaterial generiert, Abweichungen aufweist. Diese Abweichungen lassen sich vermutlich auf die weitergehende Verarbeitung (zugefügter Hall) zurückführen. Im Vergleich zu den anderen Störgeräuschen sind sie jedoch geringer und korrespondieren zu den perceptiven Ergebnissen. Die Diskriminationsfunktion des FBE im Störgeräusch DRF erzielte mit 0,6 dB SNR den höchsten  $L_{50}$  und mit 9,0 %-Punkten pro dB SNR den steilsten Verlauf. Döring und Hamacher (1992) erreichten eine ähnliche Steigung (9,6 %-Punkte pro dB SNR), während der  $L_{50}$  bei -0,1 dB SNR erreicht wurde. Die Unterschiede im SNR könnten darauf beruhen, dass die Einsilber bei Döring und Hamacher (1992) nicht nur einmal, sondern als Dreinsilber (dreimal hintereinander) präsentiert wurden. Allerdings könnten auch noch weitere Unterschiede in der Messdurchführung zu anderen Messergebnissen geführt haben. Wagener und Brand (2005) beschrieben z. B. einen Unterschied im Sprachverstehen bei der Präsentation von kontinuierlichem und zwischen den Testitern unterbrochenem Störgeräusch von 1,4 dB im Oldenburger Satztest. Für die Untersuchung von Döring und Hamacher (1992) ist nicht bekannt, ob ein kontinuierliches oder (wie hier) ein unterbrochenes Störgeräusch verwendet wurde. Mallinger (2011) verwendete u.a. wie in dieser Untersuchung eine einmalige Präsentation jedes Wortes, aber ein kontinuierliches Störgeräusch. Bei einem SNR von -2,0 dB wurde im CCITT-Rauschen ein Sprachverstehen von 47,8 % erreicht. Mit den hier erhobenen Daten würde ein solches Sprachverstehen bei einem SNR von -0,3 dB (abgeschätzt anhand des  $L_{50}$  und  $s_{50}$  für das CCITT) erreicht werden. Der Unterschied von 1,7 dB zu dem Ergebnis von Mallinger (2011) ist etwas größer als die Differenz von 1,4 dB zwischen kontinuierlichem und unterbrochenem Rauschen von Wagener und Brand (2005). Möglicherweise lässt sich die verbleibende Differenz durch die unterschiedlichen Messkonditionen oder die Messungenauigkeit erklären. Mallinger (2011) präsentierte alle 400 Wörter randomisiert und nicht nach DIN 45621-1 (1995). Weiterhin umfasste die Auswertung nur Daten von zehn und nicht wie in dieser Studie von 40 Probanden.

### Reimtest nach von Wallenberg und Kollmeier

Die Diskriminationsfunktion des WAKO im ERW-Rauschen zeigte in Analogie zum FBE im DRF-Rauschen im Vergleich zu den anderen Störgeräuschen den schlechtesten  $L_{50}$ . Einerseits wurde dies aufgrund der Generierung des Rauschens aus dem Sprachmaterial erwartet. Andererseits zeigt Abbildung 1 große Unterschiede für die Leistungsdichtespektren von WAKO und ERW im hohen Frequenzbereich. Diese Abweichungen sind erstaunlich und schlecht nachvollziehbar. Das von Müller (1992) publizierte Leistungsdichtespektrum des Rauschens im WAKO weist einen ähnlichen Frequenzgang wie das hier verwendete ERW auf. Deshalb liegt die Vermutung nahe, dass das Sprachmaterial nachträglich hochpassgefiltert wurde. Die

Unterschiede im  $L_{50}$  für die verschiedenen Störgeräusche scheinen somit durch den Grad der Verdeckung bei tiefen bis mittleren Frequenzen bestimmt zu sein.

Im Vergleich mit den Literaturdaten entspricht der  $L_{50}$  im ERW den Ergebnissen von Holube (1993), jedoch nicht denjenigen von Müller (1992). Ein Grund dafür könnte die unterschiedliche Kalibrierung sein, die bei Müller (1992) auf Spitzenpegeln beruhte. Winkler und Holube (2016) zeigten, dass das Sprachsignal bei der Kalibrierung auf Spitzenpegeln beim FBE ca. 6,5 dB geringer ist als bei der Kalibrierung auf den mittleren Langzeitpegel (sog. root-mean-square Pegel: RMS). Weitere Unterschiede zu dieser Studie sind das Alter (23 – 50 Jahre) und das Hörvermögen (beschrieben als „klinisch unauffällig“) des Probandenkollektivs von Müller (1992).

### Vergleich der Sprachtests

Der beobachtete Unterschied zwischen dem WAKO und dem FBE von ca. 12,5 dB bei gleichem Sprachverstehen kann, wie beschrieben, teilweise auf die abweichende Kalibrierung zurückgeführt werden. Der gemessene Unterschied ist etwas geringer als bei den Ruhemessungen von Sukowski et al. (2009, 2010). Sie ermittelten ohne die Berücksichtigung der Kalibrierungsunterschiede für den WAKO einen um ca. 15 dB geringeren Schalldruckpegel bei gleichem Sprachverstehen als für den FBE. Wird die Spitzenpegel-Kalibrierung beim FBE in die RMS-Pegel-Kalibrierung transformiert, dann verbleibt noch ein Unterschied von ca. 6 dB zwischen den beiden Sprachtestverfahren für die Messungen im Störgeräusch.

Mögliche Erklärungen für die Differenzen im Sprachverstehen sind das unterschiedliche Testverfahren und die geringere Verdeckungswirkung der Störgeräusche im hohen Frequenzbereich beim WAKO im Gegensatz zum FBE bei gleichen RMS (siehe Abbildung 1). Der WAKO wurde mit einem geschlossenen Antwortformat durchgeführt, sodass die Probanden möglicherweise häufiger ihnen unbekannte oder ungeläufige Wörter auswählten als beim offenen Antwortformat im FBE. Weiterhin gibt es im FBE Wörter, die im heutigen Sprachgebrauch seltener oder auch gar nicht mehr verwendet werden (Steffens 2016, Hoppe 2014). Darüber hinaus hat die Anzahl an lexikalisch ähnlichen Wörtern im Wortschatz einen Einfluss auf das Sprachverstehen (Winkler et al. 2017). Der FBE enthält außerdem Wörter, die bei den Probanden möglicherweise mit Emotionen verbunden sind (z. B. „Sau“ oder „Tod“).

### Modellierung des Sprachverstehens mittels SII

Die spektralen und kalibrierbedingten Unterschiede zwischen den Störgeräuschen und den Sprachsignalen flossen in die Modellierung des Sprachverstehens mit dem SII ein. Basierend auf dem SNR wurde die spektrale Maskierung durch die Störgeräusche auf das Sprachsignal für die verschiedenen Gewichtungsfunktionen bestimmt. Zusätzlich wurde das Sprachmaterial entsprechend dem SNR verändert und dadurch die Maskierung erhöht oder verringert.

Unabhängig von der Gewichtungsfunktion und vom Sprachtest war die mittlere quadratische Abweichung der modellierten SII-Werte von der Regressionsgeraden sehr gering. Zu beachten ist, dass die geringste mittlere quadratische Abweichung nicht für die Gewichtungsfunktion für englische Einsilber, sondern für „short passages“, die eigentlich kontinuierlich gesprochene Sprache (wie z. B. einen Vortrag oder eine Rede) repräsentieren sollte, erreicht wurde. Die Gewichtung

von „short passages“ ist im tiefen Frequenzbereich höher und im mittleren Frequenzbereich niedriger als bei den anderen Funktionen (siehe Abbildung 2). Mit dieser Gewichtungsfunktion scheint für deutsche Einsilber eine bessere Modellierung erreichbar zu sein als mit den anderen Gewichtungsfunktionen. Womöglich könnte jedoch eine spezielle, noch zu ermittelnde Funktion für den FBE das Sprachverstehen im Störgeräusch besser modellieren.

Der berechnete SII-Wert war für das CCITT-Rauschen und das Verdeckungsgeräusch nach DIN 60645 geringer als für die anderen Störgeräusche. Die Leistungsdichtespektren für CCITT und DIN 60645 wiesen im mittleren Frequenzbereich bei gleichem RMS höhere Pegel auf als ERW, DRF, ICRA1 und die Sprachsignale (negative Differenz, siehe Abbildung 1). Für den FBE widersprach dieses Modellergebnis aber den perzeptiven Ergebnissen im Sprachverstehen. Die schwierigste Situation für die Probanden war nicht die Messung im CCITT, sondern im DRF. Der SII unterschätzte also das FBE-Sprachverstehen im CCITT und überschätzte es im DRF. Das lässt darauf schließen, dass sich die perzeptiven Unterschiede im Sprachverstehen des FBE nicht allein durch den SII modellieren und damit auf spektrale Effekte der Störgeräusche zurückführen lassen. Scheinbar beeinflussen zusätzliche Faktoren, die nicht im SII enthalten sind, das Sprachverstehen. Beim WAKO hingegen konnte die Modellierung mit dem SII den Einfluss der Störgeräusche auf das Sprachverstehen nachbilden.

Ein Vergleich des  $SII_{50}$  zeigt, dass nur die Werte für den FBE in einem ähnlichen Wertebereich wie die Richtwerte nach Magnusson (1996) liegen. Für den WAKO ist der  $SII_{50}$  wesentlich geringer. Dies ist dadurch bedingt, dass beim WAKO aufgrund der bereits diskutierten Unterschiede im Testverfahren niedrigere SNR-Werte verwendet wurden, bei denen eine höhere Maskierung durch das Störgeräusch als beim FBE vorlag. Diese Unterschiede im SII zwischen FBE und WAKO sind jedoch für die praktische Anwendung der Sprachtestverfahren nicht von Bedeutung. Der SII diente in dieser Untersuchung lediglich als relatives Maß zur Analyse der spektralen Effekte der Störgeräusche auf das Sprachverstehen im  $L_{50}$  und nicht zum absoluten Vergleich der beiden Sprachteste.

### Hilfsmittelrichtlinie

Ausgangspunkt dieser Studie waren die fehlenden Angaben zum Störgeräusch in der Hilfsmittelrichtlinie (2017) für Messungen mit dem FBE im Störgeräusch. Die Ergebnisse legen nahe, dass in der Hilfsmittelrichtlinie ein Störgeräusch benannt, zumindest jedoch die Angabe des verwendeten Störgeräusches gefordert werden sollte, um Messergebnisse reproduzieren bzw. vergleichen zu können.

Zu beachten ist aber, dass die vorliegenden Ergebnisse nicht als Referenzwerte für den FBE im Störgeräusch bei der Anwendung der Hilfsmittelrichtlinie gelten können. Zur Überprüfung von Hörgeräteversorgungen sollte aufgrund des Einschwingverhaltens der Hörgeräte ein kontinuierliches Störgeräusch verwendet werden. Dies führt vermutlich durch eine Adaptation an das Störgeräusch zu einem besseren Sprachverstehen (Wagener und Brand 2005). Außerdem lassen die Messdaten keine Aussage über die perzeptiven Testlistenäquivalenz des FBE im Störgeräusch und damit zur Messgenauigkeit zu. In dieser Untersuchung wurden die vier nach Baljic et al. (2016) auffälligen Testlisten nicht verwendet. Ob diese oder andere Testlisten im Störgeräusch perzeptiv auffällig sind, muss durch weitere Untersuchungen nach den Vorgaben der DIN EN ISO 8253-3 (2012) geklärt

werden. Dafür sollten Messungen mit allen Listen des FBE im kontinuierlichen Störgeräusch erfolgen.

## Danksagung

Die Autoren danken Kristin Sprenger, Maximilian Hehl und Florian Schmitt für die Unterstützung bei der Erhebung der Messdaten, allen Probanden für die Teilnahme an der Studie und Daniel Berg (HörTech gGmbH) für die technische Unterstützung. Die Korrektur der englischsprachigen Textanteile erfolgte durch [www.stels-ol.de](http://www.stels-ol.de). Dieses Projekt wurde gefördert durch das Promotionsprogramm Jade2Pro der Jade Hochschule.

## Literatur

- ANSI S3.5 (1997) Methods for calculation of the Speech Intelligibility Index (SII). New York: Acoustical Society of America
- Baljic I, Winkler A, Schmidt T und Holube I (2016) Untersuchungen zur perzeptiven Äquivalenz der Testlisten im Freiburger Einsilbertest. In: HNO 64 (8), 572–583
- Beutelmann R und Brand T (2006) Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners. In: J. Acoust. Soc. Am. 120 (1), S. 331–342
- Brand T und Kollmeier B (2002). Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. In: J. Acoust. Soc. Am. 111 (6), S. 2801–2810
- Brand T und Wagener KC (2005). Wie lässt sich die maximale Verständlichkeit optimal bestimmen? In: 8. Jahrestagung Deutsche Gesellschaft für Audiologie: 24.-26.02.2005 Göttingen; Tagungs-CD. Oldenburg: Dt. Ges. für Audiologie
- Brinkmann K (1974). Die Neuaufnahme der „Wörter für Gehörprüfung mit Sprache“. Zeitschrift für Hörgeräteakustik (13), S. 14–40
- DIN 45621-1 (1995) Sprache für Gehörprüfung. Teil 1: Ein- und mehrsilbige Wörter. Beuth Verlag, Berlin
- DIN 45626-1 (1995) Tonträger mit Sprache für Gehörprüfung, Teil 1: Tonträger mit Wörtern nach DIN 45621-1. Beuth Verlag, Berlin
- DIN EN 60645-2 (1997) Audiometer Teil 2: Geräte für die Sprachaudiometrie. Beuth Verlag, Berlin
- DIN EN ISO 8253-1 (2011) Akustik - Audiometrische Prüfverfahren - Teil 1: Grundlegende Verfahren der Luft- und Knochenleitungs-Schwellenaudiometrie mit reinen Tönen. Beuth Verlag, Berlin
- DIN EN ISO 8253-3 (2012) Akustik - Audiometrische Prüfverfahren - Teil 3: Sprachaudiometrie. Beuth Verlag, Berlin
- Döring W und Hamacher V (1992) Neue Sprachverständlichkeitstests in der Klinik: Aachener Logatomtest und Dreinsilber im Störschall. In: Kollmeier B (Hg.) Moderne Verfahren der Sprachaudiometrie, Median-Verlag von Killisch-Horn GmbH, Heidelberg, S. 137–168
- Dreschler WA, Verschuur H, Ludvigsen C und Westermann S (2001) ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. International Collegium for Rehabilitative Audiology. In: Audiology 40 (3), S. 148–157
- Festen JM und Plomp R (1990) Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. In: J. Acoust. Soc. Am. 88 (4), S. 1725–1736
- Hahlbrock KH (1953) Über Sprachaudiometrie und neue Wörterteste. In: Archiv Ohr- usw. Heilk. u. Z. Hals- usw. Heilk. (162), S. 394–431
- Hilfsmittelrichtlinie (2017) Richtlinie des gemeinsamen Bundesausschusses über die Verordnung von Hilfsmitteln in der vertragsärztlichen Versorgung. [https://www.g-ba.de/downloads/62-492-1352/Hilfsm-RL\\_2016-11-24\\_iK-2017-02-17.pdf](https://www.g-ba.de/downloads/62-492-1352/Hilfsm-RL_2016-11-24_iK-2017-02-17.pdf) (Stand 02.03.2017)
- Holube I (1993) Experimente und Modellvorstellungen zur Psychophysik und zum Sprachverstehen bei Normal- und Schwerhörigen. Dissertation: Georg-August-Universität Göttingen
- Holube I und Kollmeier B (1996) Speech intelligibility prediction in hearing-impaired listeners based on a psychoacoustically motivated perception model. In: J. Acoust. Soc. Am. 100 (3), S. 1703–1716
- Holube I, Winkler A und Nolte-Holube R (2018) Modellierung der Reliabilität des Freiburger Einsilbertests in Ruhe mit der verallgemeinerten Binomialverteilung: Hat der Freiburger Einsilbertest 29 Wörter pro Liste? In: Zeitschrift für Audiologie 57 (1), S. 6–17
- Holube I, Winkler A und Simmer U (2017) CCITT, ICRA & Co.: Kurzmittelung. In: Zeitschrift für Audiologie 56 (2), S. 58–59
- Hoppe U (2014) Freiburger Sprachverständlichkeitstest – überflüssig oder notwendig? In: Zeitschrift für Hörgeräteakustik (7), S. 6–9
- HörTech gGmbH. <http://www.hoertech.de> (Stand 06.10.2015)
- ITU Recommendation G.227 (1993) Conventional telephone signal. <https://www.itu.int/rec/T-REC-G.227-199811-1/en> (Stand 12.04.2016)

- Lazarus H, Sust CA, Steckel R, Kulka M und Kurtz P (2007) Akustische Grundlagen sprachlicher Kommunikation (German Edition). Springer, Dordrecht
- Magnusson L (1996) Speech Intelligibility Index: transfer functions and speech spectra for two Swedish speech recognition tests. In: Scand Audiol 25 (1), S. 59–67
- Mallinger E (2011) Trainingseffekte und Listenäquivalenz des Freiburger Einsilbertests im Störschall. Dissertation: Friedrich-Alexander-Universität Erlangen-Nürnberg <https://d-nb.info/1015782426/34> (Stand 14.05.2018)
- Müller C (1992) Perzeptive Analyse und Weiterentwicklung eines Reimtestverfahrens für die Sprachaudiometrie. Dissertation: Georg-August-Universität Göttingen
- Rhebergen KS, Lyzenga J, Dreschler WA und Festen JM (2010) Modeling speech intelligibility in quiet and noise in listeners with normal and impaired hearing. In: J. Acoust. Soc. Am. 127 (3), S. 1570–1583
- Steffens T (2016) Die Verwendungshäufigkeit der Freiburger Einsilber in der kontemporären deutschen Sprache - Sind die Testwörter noch zeitgemäß? In: HNO 64 (8), S. 549–556
- Sukowski H, Brand T, Wagener KC und Kollmeier B (2009) Untersuchung zur Vergleichbarkeit des Freiburger Sprachtests mit dem Göttinger Satztest und dem Einsilber-Reimtest nach von Wallenberg und Kollmeier. In: HNO 57 (3), S. 239–250
- Sukowski H, Brand T, Wagener KC und Kollmeier B (2010) Vergleich des Göttinger Satztests und des Einsilber-Reimtests nach von Wallenberg und Kollmeier mit dem Freiburger Sprachtest: Untersuchung bei einem klinisch repräsentativen Probandenkollektiv. In: HNO 58 (6), S. 597–604
- von Wallenberg E-L und Kollmeier B (1989) Sprachverständlichkeitsmessungen für die Audiologie mit einem Reimtest in deutscher Sprache: Erstellung und Evaluation von Testlisten. In: Audiologische Akustik 28 (2), S. 50–56
- Wagener KC (2004) Factors influencing sentence intelligibility in noise. Dissertation: Carl von Ossietzky Universität Oldenburg
- Wagener KC und Brand T (2005) Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: Influence of measurement procedure and masking parameters La inteligibilidad de frases en silencio para sujetos con audición normal y con hipoacusia: la influencia del procedimiento de medición y de los parámetros de enmascaramiento. In: Int J Audiol 44 (3), S. 144–156
- Wagener KC, Kühnel V und Kollmeier B (1999) Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests. In: Zeitschrift für Audiologie 38 (1), S. 4–15
- Wesselkamp M, Kliem K und Kollmeier B (1992) Erstellung eines optimierten Satztests in deutscher Sprache. In: Kollmeier B (Hg.) Moderne Verfahren der Sprachaudiometrie, Median-Verlag von Killisch-Horn GmbH, Heidelberg, S. 330–343
- Westra Elektroakustik GmbH. CD 8 - Dreinsilber-Test Freiburger Wörter mit speziellem Störgeräusch. <https://www.westra.de/hoerdiagnostik/audiologische-tests-fuer-erwachsene/cd-8-dreinsilber-test-freiburger-woerter-im-stoergeraesch/> (Stand 23.10.2017)
- Winkler A, Carroll R und Holube I (2017) Analyse von Nachbarschaftsparametern anhand der Freiburger Einsilber. In: 20. Jahrestagung Deutsche Gesellschaft für Audiologie: Aalen, 22.-25.02.2017, Tagungs-CD
- Winkler A und Holube I (2014) Was wissen wir über den Freiburger Sprachtest? In: Zeitschrift für Audiologie 53 (4), S. 146–154
- Winkler A und Holube I (2016) Der Freiburger Einsilbertest und die Norm DIN EN ISO 8253-3: Technische Analyse. In: Zeitschrift für Audiologie 55 (3), S. 106–113

*Alexandra Winkler beendete im Jahr 2000 erfolgreich ihre Ausbildung zur Hörgeräteakustikerin und studierte Hörtechnik und Audiologie an der Jade Hochschule und der Carl von Ossietzky Universität in Oldenburg. Abgeschlossen hat sie diese Studiengänge sowohl mit dem Diplom (2005) als auch mit dem Master (2007). Anschließend war sie als wissenschaftliche Mitarbeiterin an der Carl von Ossietzky Universität in Oldenburg tätig. Am Institut für Hörtechnik und Audiologie der Jade Hochschule arbeitet sie seit 2009 als wissenschaftliche Mitarbeiterin mit den Forschungsschwerpunkten Hörgerätesstechnik und audiologische Studien. Im Dezember 2014 begann sie ihre Promotion mit dem Thema „Validierung von Hörgeräteanpassungen mit Sprachtestverfahren im Labor und im Alltag“ im Rahmen des Jade2Pro Programms der Jade Hochschule in Oldenburg.*



(Fotos: Jade Hochschule)

*Prof. Dr. Inga Holube studierte ab 1984 Physik und promovierte 1993 unter Anleitung von Prof. Dr. Dr. Birger Kollmeier am III. Physikalischen Institut der Universität Göttingen. Zwischen 1993 und 1995 war sie wissenschaftliche Mitarbeiterin bei Kollmeier an der Universität Oldenburg und ein Jahr lang in den Niederlanden, USA und England im Rahmen eines Postdoktoranden-Stipendiums der Deutschen Forschungsgemeinschaft (DFG) tätig. Von 1995 bis 2001 leitete sie die Abteilung Audiologische Grundlagen und Systemtechnik bei Siemens Audiologische Technik GmbH. Seit September 2001 ist sie Professorin im Studiengang Hörtechnik und Audiologie an der Fachhochschule Oldenburg/Ostfriesland/Wilhelmshaven, jetzt Jade Hochschule.*

