

# Zusammenhang zwischen Sprachverständlichkeitsschwelle in fluktuierenden Störgeräuschen und nicht-auditorischen Faktoren bei Normalhörenden.

*Ralf Meyer, Thomas Brand, Birger Kollmeier*

Carl-von-Ossietzky Universität Oldenburg, Medizinische Physik, Oldenburg

## Einleitung

Die Messung von Sprachverständlichkeitsschwellen ist in der Diagnostik ein wichtiges Instrument bei der Charakterisierung des Hörvermögens und der Sprachverarbeitung. Die Messung von Sprachverständlichkeitsschwellen in fluktuierenden Störgeräuschen bietet den Vorteil, dass sich in den Ergebnissen solcher Messungen große Unterschiede zwischen verschiedenen Hörverlusten zeigen, die bei stationären Störgeräuschen nicht in dieser Form vorhanden sind (z. B. Wagener et al. 2006). Die Vorhersage der Schwellen bietet zum einen die diagnostische Möglichkeit der Erkennung auffälliger Befunde, zum anderen ist es möglich, anhand des verwendeten Modells Mechanismen der Sprachverarbeitung zu verstehen und zu untersuchen.

Die Vorhersage der Schwellen basiert bei den aktuellen Modellen (ANSI S3.5-1997; Brand et al., 2003; Rhebergen et al. 2005) auf dem Ton-Audiogramm, deshalb ergeben sich bei normalhörenden Versuchspersonen ( mit sehr ähnlichen Audiogrammen) auch sehr ähnliche Vorhersagen für die Sprachverständlichkeitsschwellen (Meyer et al., 2007). Die beobachteten Schwellen zeigen jedoch eine deutliche Streuung. Aus diesem Grund sollten weitere auch nicht-auditorische Faktoren (wie z. B. die linguistische Kompetenz) bei der Vorhersage von Sprachverständlichkeitsschwellen berücksichtigt werden. Das Ziel dieser Studie war es deshalb, den Zusammenhang zwischen den gemessenen Sprachverständlichkeitsschwellen in fluktuierenden Störgeräuschen und anderen nicht-auditorischen Faktoren zu untersuchen. Dabei wurden als nicht-auditorische Faktoren, die Ergebnisse aus dem „Lexical Decision“-Test, sowie die Textverständlichkeitsschwelle untersucht.

## Experimente

Für zunächst fünf normalhörende Versuchspersonen wurden die Sprachverständlichkeitsschwellen in verschiedenen fluktuierenden und stationären Störgeräuschen mit dem Oldenburger Satztest (Wagener et al., 1999) gemessen. Dabei wurden die folgenden Störgeräusche verwendet:

- ICRA1 (Dreschler et al. 2001, stationäres Störgeräusch, mit dem Langzeitspektrum von Sprache),
- ICRA1 rechteckmoduliert (4Hz),
- ICRA5-250 Störgeräusch (Wagener et al., Simulation von einem männlichen „Stör-Sprecher“),
- ISTS (Sprachähnliche Modulationen, weiblicher Sprecher),
- Sätze des Göttinger Satztest (zusätzliche Maskierung durch Information).

Die Störgeräusche unterscheiden sich in ihrer Sprachähnlichkeit. Während das ICRA1-Rauschen nur das Langzeitspektrum von Sprache nachbildet, zeigt das rechteckmodulierte ICRA1-Rauschen für Sprache untypische und das ICRA5-250 für Sprache typische Modulationen. Das ISTS Störgeräusch weist eine große Sprachähnlichkeit auf ohne jedoch fließende Sprache zu sein. Bei den Sätzen des Göttinger Satztestes handelt es sich um eigentliche Sprache. Für jedes Störgeräusch wurden jeweils monaural (besseres Ohr) mindestens zwei Schwellen pro Versuchsperson gemessen und dann über diese Schwellen gemittelt. Die Versuchspersonen hatten vor jeder Sitzung zwei Trainingslisten mit zufällig aus dem obigen Pool ausgewählten Störgeräuschen zu absolvieren.

Beim „Lexical Decision“-Test (Hällgren et al. 2007) geht es um die Fähigkeit, deutsche Worte von zufälligen Buchstabenkombinationen in möglichst kurzer Zeit zu unterscheiden. Die Versuchspersonen sehen auf einem Bildschirm das Test-“Wort“ und müssen dann durch Tastendruck entscheiden, ob ein deutsche Wort dargestellt wird oder nicht. Dabei wurde die Trefferquote, sowie die Antwortzeit gemessen. Vor der eigentlichen Messung erfolgte eine Übungsliste mit 12 Worten. Das Ergebnis wird aus der Mittellung von zwei aufeinander folgenden

Wortlisten mit jeweils 50 Wörtern errechnet. Der „Lexical Decision“-Test stammt aus dem „Auditory Profile“-Inventar des EU-Projektes „Hearcom“.

Die Messung der Textverständlichkeitsschwelle erfolgt mit dem so genannten TRT-Test (Zekveld et al. 2007). Dabei sehen die Versuchspersonen einen Satz aus dem Oldenburger Satztest auf einem Bildschirm der Satz wird dabei von einem Muster aus schwarzen Formen überlagert. Das Muster überdeckt einen definierten Prozentsatz der Fläche, auf der der Satz dargestellt wird. Der Prozentsatz der Überdeckung wird adaptiv (vergleichbar mit dem adaptiven Oldenburger Satztest) angepasst. Das Ergebnis ist der Prozentsatz der Überdeckung für 50% Verständlichkeit beim Lesen des Satzes. Der Test wurde als Analogon für den Oldenburger Satztest entwickelt. Die Textverständlichkeitsschwelle wurde für drei verschiedene Muster ermittelt: quadratische zufällig verteilte Punkte (Punktgröße größer als Schriftbreite), periodische Balken an festen Positionen (die Dicke der Balken wurde verändert) und sprachähnlich verteilte Balken (Verteilung der Balken angeglichen an die Einhüllende des ICRA5-250-Rauschens). Die Punkte sind einem stationären „Störgeräusch“ und die Balken einem fluktuierendem „Störgeräusch“ nachempfunden.

Zusätzlich wurden ein Ton-Audiogramm und eine Kategoriale Lautheitsskalierung bei allen Versuchspersonen gemessen. Die Lautheitsskalierung wurde bei 500Hz, 1000Hz, 2000Hz, 3000Hz und 4000Hz, sowie breitbandig mit dem Rauschen des Oldenburger Satztest durchgeführt.

## Ergebnisse

Die Abbildungen 1 bis 4 zeigen die Ergebnisse für die verschiedenen Tests. Abbildung 1 zeigt die Ergebnisse für den Oldenburger Satztest für die verschiedenen Störgeräusche. Für alle Versuchspersonen und alle Störgeräusche ergeben sich negative Signal-Rausch-Verhältnisse (S/N) für die 50%-Schwellen. Für das ICRA1-Rauschen ergeben sich die höchsten Schwellen (zwischen -6,35 und -8,75 dB S/N). Die niedrigsten Schwellen ergeben sich für das ISTS-Rauschen (zwischen -29,5 und -24,9 dB S/N). Danach folgt das ICRA5-250-Rauschen mit Werten zwischen -23,5 und -19,25 dB S/N. Das rechteck-modulierte ICRA1-Rauschen und die Sätze des Göttinger Satztest liegen bei fast allen Versuchspersonen fast gleich auf zwischen -18 und -22 dB S/N. Abbildung 2 zeigt die Ergebnisse für den TRT-Test. Der Grad der Überdeckung schwankt für alle Versuchspersonen und alle Störmuster zwischen 61% und 76%. Für fast alle Versuchspersonen ist das sprachähnliche Störmuster am schwersten und ergibt die niedrigsten Prozentzahlen. Die periodischen Balken und die Punkte liefern eine höhere Überdeckung bei 50% Lesbarkeit.

Die Abbildungen 3 und 4 zeigen die Ergebnisse für den „Lexical Decision“-Test. Abbildung 3 zeigt die Trefferquote der einzelnen Versuchspersonen. Die Quote liegt zwischen 94% und 100%, es zeigen sich also nur geringe interindividuelle Unterschiede. Die Antwortzeiten sind in Abbildung 4 zu sehen, sie liegen zwischen 540 ms und 830 ms, hier zeigen sich größere interindividuelle Schwankungen als bei der Trefferquote.

Die Abbildungen 5 und 6 zeigen den Zusammenhang zwischen den Sprachverständlichkeitsschwellen für die verschiedenen Störgeräusche und den Ergebnissen des TRT-Tests (Abb. 5 für periodische Balken und Abb. 6 für sprachähnliche Balken). Aufgrund der geringen Versuchspersonenanzahl sind keine nennenswerten Korrelationen zwischen Sprachverständlichkeitsschwelle und Texterkennungsschwelle zu erkennen.

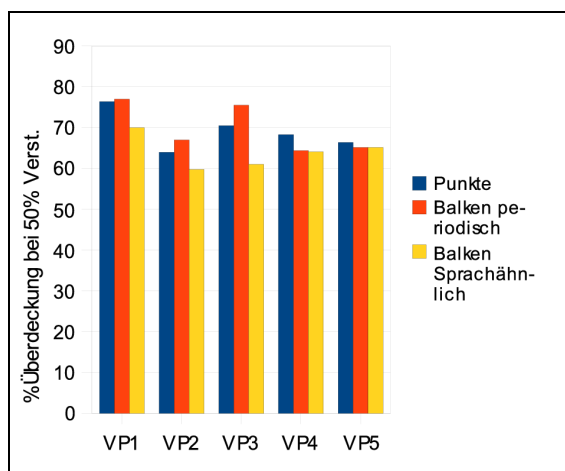


Abbildung 1: Textverständlichkeitsschwelle in Prozent der überdeckten Fläche, für alle fünf Versuchspersonen für alle Störmuster (codiert durch die unterschiedlichen Farben).

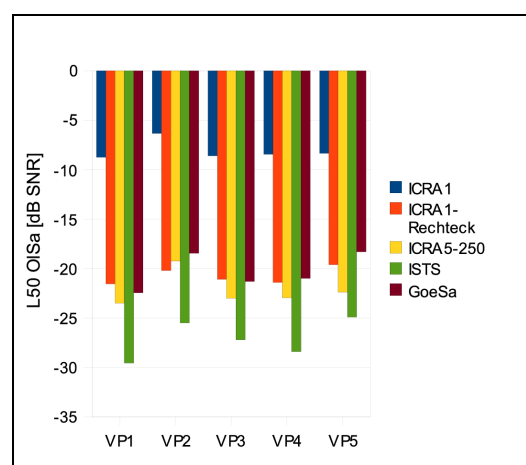


Abbildung 2: Sprachverständlichkeitsschwellen in dB SNR für alle fünf Versuchspersonen für alle Störgeräusche (codiert durch die unterschiedlichen Farben).

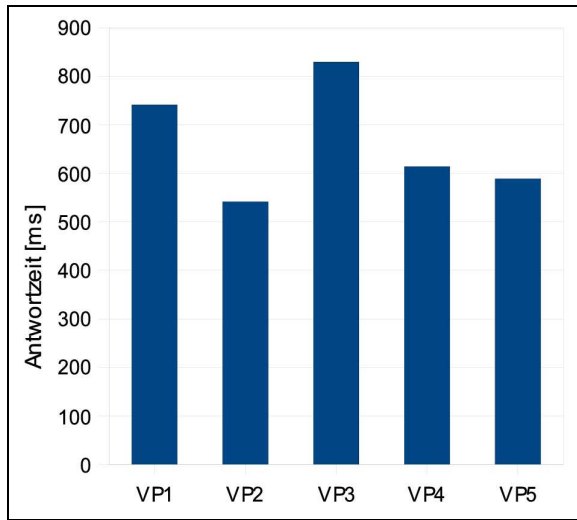


Abbildung 3: Antwortzeit in Millisekunden für alle fünf Versuchspersonen beim "Lexical Decision"-Test

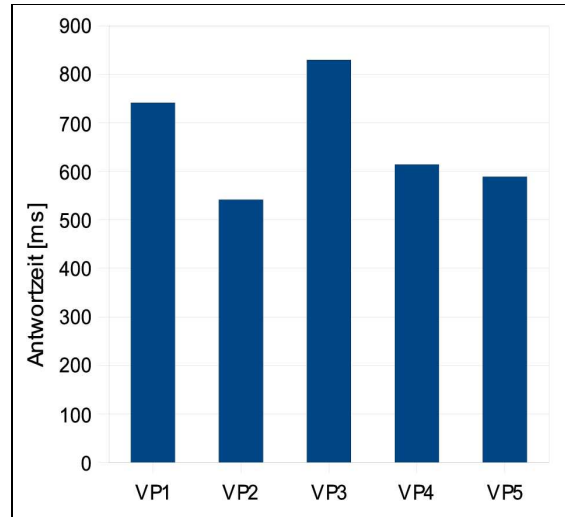


Abbildung 4: Trefferquote in Prozent für alle fünf Versuchspersonen beim "Lexical Decision"-Test

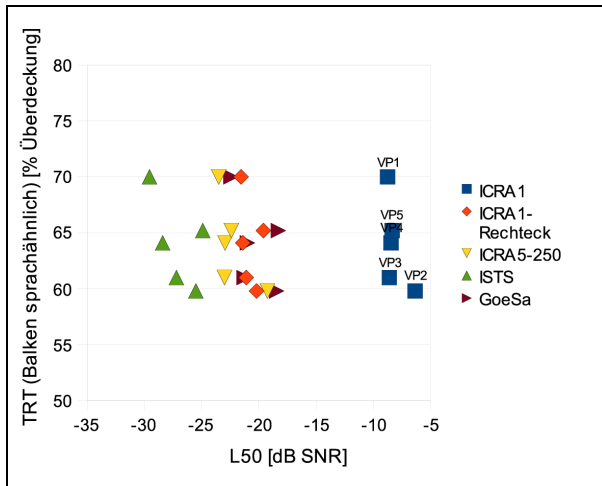


Abbildung 5: Zusammenhang zwischen L50 in dB SNR (x-Achse) und TRT in % Überdeckung bei sprachähnlichen Balken (y-Achse). Für alle fünf Versuchspersonen. Die unterschiedlichen Symbole und Farben zeigen die verschiedenen Störgeräusche.

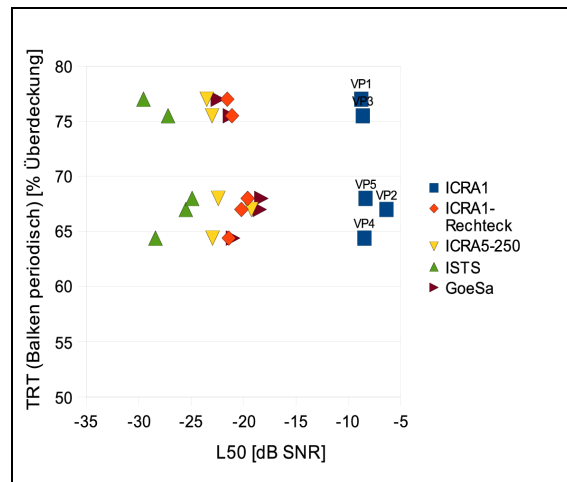


Abbildung 6: Zusammenhang zwischen L50 in dB SNR (x-Achse) und TRT in % Überdeckung bei periodischen Balken (y-Achse). Für alle fünf Versuchspersonen. Die unterschiedlichen Symbole und Farben zeigen die verschiedenen Störgeräusche.

## Diskussion und Zusammenfassung

Ziel der Studie war eine Unterscheidung zwischen sensorischen und kognitiven Komponenten bei der Sprachperzeption mit fluktuierendem Störschall, da bislang die Modellierung mit rein sensorischen Faktoren die interindividuelle Streuung nicht erklären kann (Meyer et al., 2007). Die Annahme ist, dass diese Differenz zwischen Messung und Modell durch kognitive Faktoren erklärt werden könnte.

Die Experimente zur Messung der Sprachverständlichkeitsschwelle bei verschiedenen fluktuierenden Störgeräuschen erbrachte für das ISTS Störgeräusch die niedrigsten Schwellen. Bei diesem Störgeräusch ist der perzeptive Unterschied zwischen Zielsatz und Störgeräusch am größten ist. Das ISTS Störgeräusch besitzt das Frequenzspektrum einer weiblichen Stimme, während die Zielsätze von einem männlichen Sprecher gesprochen werden. Für das ICRA5-250 Störgeräusch und die Sätze des Göttinger Satztests wurden höhere Schwellen gemessen. Diese Störgeräusche haben das Frequenzspektrum einer männlichen Stimme, bzw. die Sätze des Göttinger Satztests sind sogar vom gleichen Sprecher gesprochen. Daher ist der perzeptive Unterschied zwischen Zielsatz und Störgeräusch nicht so groß wie beim ISTS Störgeräusch.

Kognitive Komponenten sollten mit dem „Lexical Decision“-Test und dem TRT-Test erfasst werden. Die Antwortzeit beim Lexical Decision Test differenziert interindividuell am besten, die Trefferquote dagegen war bei allen Versuchspersonen gleichermaßen hoch. Der TRT-Test hat kaum Unterschiede zwischen Störmustern

und Probanden gezeigt. Die angenommene Analogie, dass die Punkte ein stationäres Störgeräusch darstellen und die Balken ein fluktuierendes, scheint nicht gegeben zu sein. Die Ergebnisse für die Punkte lagen nicht deutlich unter den Ergebnissen für die Balken.

Betrachtet man die Korrelationen zwischen den auditorischen und den kognitiven Tests, so sind bis jetzt keine nennenswerten interindividuellen Korrelationen zwischen TRT-Test und Satztest zu erkennen. Da die Zahl der Versuchspersonen und die interindividuelle Variation der Sprachverständlichkeitsschwellen und der kognitiven Tests zu gering ist, sind weitere Messungen mit normal- und schwerhörigen Versuchspersonen notwendig, um eine konkrete Aussage über den Zusammenhang der beiden Komponenten zu treffen. Bei der Auswertung des „Lexical Decision“-Test ist eine Kombination von Antwortzeit und Score sinnvoll. Weiterhin ist es die Integration der nicht-auditorischen, kognitiven Faktoren in die Vorhersagemodelle (Meyer et al., 2007) geplant.

## Danksagung

Wir danken der Audiologie Initiative Niedersachsen für die Finanzierung dieser Studie, sowie dem EU-Projekt „Hearcom“ für die Unterstützung.

## Literatur

- ANSI (1997)*. „Methods for Calculation of the Speech Intelligibility Index“, American National Standard S3.5-1997
- Brand, T. (2002)*. „Vorhersage der Sprachverständlichkeit in Ruhe und im Störgeräusch aufgrund des Reintonaudiogramms“, DGA 2002
- Dreschler, W. A. (2001)*. „ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. International Collegium for Rehabilitative Audiology.“, Audiology 2001, 40, 148-157
- Hällgren, M.; Larsby, B.; Lyxell, B. & Arlinger, S. (2007)*. „Evaluation of a cognitive test battery in young and elderly normal-hearing and hearing-impaired persons.“, J. Am. Acad. Audiol., 2001, 12, 357-370
- Rhebergen, K. & Versfeld, N. (2005)*. „A Speech Intelligibility Index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners.“ J Acoust Soc Am, 117, 2181-92
- Meyer, R., Brand T., Kollmeier B. (2007)*. „Predicting speech intelligibility in fluctuating noise.“, DGA 2007
- Wagener K., Brand T., Kühnel V. und Kollmeier B. (1999)*. „Entwicklung und Evaluation eines Satztestes für die deutsche Sprache I-III: Design, Optimierung und Evaluation des Oldenburger Satztestes“, Zeitschrift für Audiologie 38(1-3)
- Wagener K., Brand T., Kollmeier B. (2006)*. „The role of silent intervals for sentence intelligibility in fluctuating noise in hearing-impaired listeners“, Int J Audiol, 45, 26-3
- Zekveld A. A., George E. L. J., Kramer S. E., Goverts S. T., Houtgast T. (2007)*. „The development of the text reception threshold test: a visual analogue of the speech reception threshold test“, J Speech Lang Hear Res, 50, 576-584