

Untersuchung der Antwortzeiten beim Freiburger Einsilbertest im Störgeräusch

Stephan Müller^{1,2}, Jürgen Tchorz², Hendrik Husstedt¹

¹Deutsches Hörgeräte Institut, Lübeck

²Fachhochschule Lübeck, Lübeck

Schlüsselwörter: Höranstrengung, Antwortzeiten, Freiburger Einsilbertest,

Einleitung

Die fortschreitende Entwicklung im Bereich der Hörgeräteversorgung hat viele neue Funktionen zur Verbesserung des Sprachverstehens hervorgebracht. Einige dieser Funktionen zielen aber auch auf eine Verringerung der Anstrengung, welche für das Sprachverstehen aufgebracht werden muss. Um die Verringerung der sogenannten Höranstrengung objektiv untersuchen zu können, sind schnelle und praxistaugliche Messverfahren wünschenswert. Messungen physiologischer Reaktionen bspw. mittels Pupillometrie (Hyönä et al., 1995) oder Elektroenzephalografie (Bernarding et al., 2017) lieferten bereits vielversprechende, jedoch aufwändige Methoden zur Erfassung der Höranstrengung. Weniger aufwändige Methoden wurden in Form von Dual-Task-Paradigmen (Sarampalis et al., 2009) oder Erfassungen der subjektiven Bewertung der Höranstrengung (Krüger et al., 2015) bereitgestellt. Auch in Form von Single-Task-Paradigmen ließen sich Änderungen der Höranstrengung etwa anhand der Antwortzeiten bei Sprachtests (Gatehouse und Gordon, 1990) nachweisen. In der alltäglichen Praxis des Hörakustikers in Deutschland stellt der Freiburger Sprachverständlichkeitstest ein Standard-Verfahren zur Erfassung des Sprachverstehens dar. In dieser Studie soll untersucht werden, ob sich die Messung der Antwortzeiten beim Freiburger Einsilbertest (FET) im Störgeräusch als Verfahren zur Untersuchung der Höranstrengung eignet. Dabei steht im Fokus, ob sich die Antwortzeiten im Bereich leisen Störgeräuschs stärker ändern als das Sprachverstehen. Des Weiteren wird Wert auf Tauglichkeit für die alltägliche Anwendung im Bereich der Hörakustik gelegt.

Materialien und Methoden

Zur Darbietung des Testmaterials und zum Festhalten der Antworten wird ein Programm in MatLab geschrieben. 18 von 20 Listen des FET werden randomisiert ausgewählt und auf einen festen Pegel von 65 dB eingestellt. CCITT-Rauschen wird zusätzlich bei verschiedenen Pegeln dargeboten. So ergeben sich die SNRs listenweise bei -10, -5, 0, 5, 10 und 40 dB. Letzterer repräsentiert dabei eine ruhige Umgebung, um die Antwortzeiten in einen Bezug zu möglichst müheloser Sprachverständlichkeit zu setzen. Die verschiedenen SNRs werden in drei Blöcke à sechs Listen eingeteilt, sodass in jedem Block jeder SNR nur einmal vorkommt. Die Reihenfolge der Listen wird randomisiert, die der Wörter bleibt unverändert. Das Rauschen sowie die Einsilber werden entsprechend der Vorgabe aus der DIN 45626-1 kalibriert.

Die Antwortzeiten bilden sich aus der Zeitdifferenz zwischen der Darbietung des Wortes und der verbalen Antwort des Probanden. Diese Differenz wird hier auf zwei verschiedene Weisen erfasst. Zum einen wird die gesamte Messung mit einem Mikrofon am Mund des Probanden aufgezeichnet, sodass die Antwortzeiten automatisiert aus dem Zeitsignal bestimmt werden können. Als Alternative dazu wird mithilfe des Messprogramms, die Zeit gemessen, die zwischen dem Drücken („Clicken“) der Schaltfläche zur Darbietung des Wortes und dem Drücken der Schaltfläche *Richtig* oder *Falsch* zur Erfassung der Antwort vergeht. Hiervon wird dann noch die Dauer des jeweiligen Einsilbers abgezogen. Diese rein Software-basierte Methode der Antwortzeit-Erfassung wird im Folgenden als Click2Click (C2C) bezeichnet. Die C2C-Antwortzeiten beinhalten sowohl die Reaktionszeit des Testleiters als auch die Dauer des Nachsprechens. Hier soll untersucht werden, ob die Differenz zwischen beiden Verfahren unabhängig vom SNR ist. Bei Ausbleiben einer Antwort oder Angaben wie „nein“ oder „nicht verstanden“ macht die Erfassung der Antwortzeit keinen Sinn, daher besteht die Option für den Testleiter, die Antwort mit einer Schaltfläche *Keine Angabe* zu bewerten. Die so erfassten Antworten fließen nicht in die Auswertung der Antwortzeiten mit ein, werden in Bezug auf das Sprachverstehen jedoch als *falsch* gewertet.

Zusätzlich zu den Antwortzeiten wird das Sprachverstehen, sowie eine subjektive Bewertung der Höranstrengung erfasst. Letztere erfolgt durch die 14-stufige Skala nach M. Krüger (2015). Die Skala beinhaltet die Stufen „mühelos“, „sehr wenig anstrengend“, „wenig anstrengend“, „mittelgradig anstrengend“, „deutlich anstrengend“, „sehr anstrengend“ und „extrem anstrengend“. Zusätzlich sind Zwischenstufen vorhanden, sowie die 14. Stufe „nur Störgeräusch“. Diese soll gewählt werden, sofern keine Sprache vernommen wurde und wird aus der Auswertung ausgeschlossen. Die Bewertung der Höranstrengung erfolgt listenweise.

40 normal hörende Probanden zwischen 20 und 30 Jahren (im Mittel 24.3 Jahre) ohne nennenswerte Erfahrung mit dem Freiburger Einsilbertest nehmen an der Studie teil. Das Hörvermögen wird am Untersuchungstag überprüft, ausgeschlossen wird bei einem Hörverlust > 20 dB bei mindestens einer Prüffrequenz. Die Einweisung der Probanden erfolgte entsprechend der beim FET üblichen. Die Probanden wussten nicht, dass die Antwortzeiten erfasst werden.

Vergleich der Click2Click-Antwortzeiten mit den automatisch erfassten Antwortzeiten

Zur Erfassung der Antwortzeiten aus der Aufnahme des Mikrofons wird die Aufnahme einer Liste jeweils pro Wort geschnitten. Beim Schnitt entspricht der Anfang dem Zeitpunkt, an dem die Darbietung des Wortes endet (aus Software bekannt), und das Ende entspricht dem Zeitpunkt, an dem die Eingabe der Antwort durch den Testleiter erfolgt ist (durch das Drücken der Schallfläche bekannt). Über eine gleitende Pegelberechnung wird mithilfe einer Schwellwertdetektion der Beginn der Antwort bestimmt und festgehalten.

Abbildung 1 zeigt die Gegenüberstellung beider Verfahren in Form eines Streudiagramms. An der Ordinate sind die automatisch erfassten Antwortzeiten aufgetragen und an der Abszisse die C2C-Antwortzeiten. Jeder Punkt repräsentiert die Antwortzeiten einer Gruppe, deren Farbe markiert den SNR. Der Pearson-Korrelationskoeffizient beträgt $r = 0.88^{***}$. Der Zusammenhang lässt sich durch die rote Ausgleichsgerade darstellen, aus deren Abschnitt mit der Abszisse sich eine konstante Verzögerung von 0.55 s durch das Nachsprechen und die Reaktionszeit des Operators ergibt. Aufgrund der hohen Korrelation zwischen beiden Größen werden im Folgenden ausschließlich die Ergebnisse durch die C2C-Antwortzeiten betrachtet. Werden die Ausreißer bei den Messdaten aus Abb. 1 genauer betrachtet, kann festgestellt werden, dass diese ausschließlich bei der automatischen Ermittlung der Antwortzeit aufgetreten sind. Dies könnte vermutlich durch weiter entwickelte Algorithmen verbessert werden. Es zeigt aber auch die Schwierigkeiten in der Praxis, die bei einer automatischen Ermittlung der Antwortzeit durch Störgeräusche auftreten können.

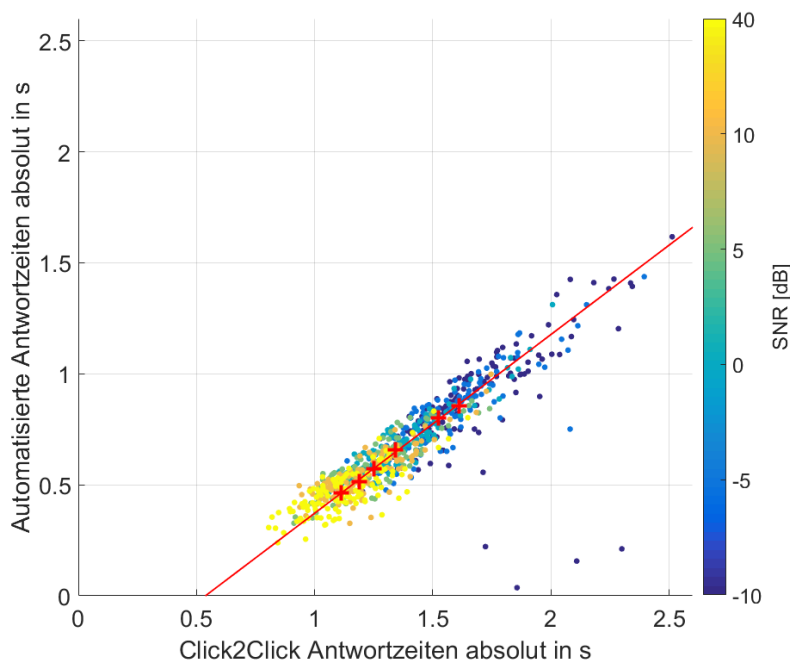


Abbildung 1: Streudiagramm von automatisch bestimmten Antwortzeiten (Ordinate) gegenüber Click-to-Click-Antwortzeiten (C2C) (Abszisse) mit Ausgleichsgerade (rot).

Ergebnisse

Die Antwortzeiten weisen in dieser Untersuchung eine hohe Streuung auf, welcher teilweise auf interindividuelle Unterschiede bei den Reaktionszeiten der Probanden zurückzuführen ist. Um die Streuung zu minimieren wird deshalb eine Normierung durchgeführt. Hierfür werden die gruppenweise gemittelten Antwortzeiten für jeden Probanden auf das Mittel der Antwortzeiten bei 40 dB SNR normiert, ausgehend von minimaler Höranstrengung in dieser Bedingung. Abb. 2 zeigt die relativen Antwortzeiten über dem SNR als Boxplots. Die Rote Kurve zeigt eine Spline-Interpolation, die als Stützstellen die Mediane der Ergebnisse verwendet. Der Verlauf zeigt, dass die Antwortzeiten entsprechend der erwarteten Höranstrengung mit steigendem SNR abfällt. Die schnellsten Antwortzeiten liegen bei 40 dB SNR. Gepaarte T-Tests zwischen den einzelnen Bedingungen weisen in jeder Kombination auf hoch signifikante Unterschiede ($p < 0.001$) hin. Die P-Werte wurden aufgrund der multiplen Vergleiche mit der Bonferroni-Korrektur angepasst.

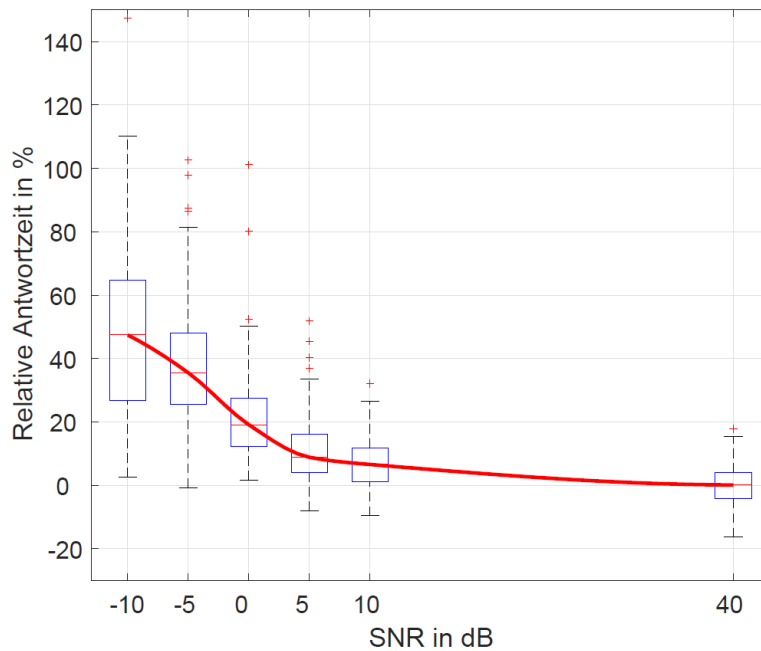


Abbildung 2: Boxplots der relativen Antwortzeiten gemittelt über alle Probanden über dem SNR. Ausreißer sind mit ‚+‘ gekennzeichnet. Der rote Funktionsverlauf wird durch einen Spline approximiert, dessen Stützstellen aus den Medianen der Messergebnisse gebildet werden.

Antworten der Probanden welche mit *keine Angabe* bewertet wurden, fließen in die Betrachtung der Antwortzeiten nicht mit ein. Dies führt dazu, dass sich die gemittelten Daten insbesondere bei negativen SNRs aus weniger Werten zusammensetzen. In Tab. 1 ist zur Verdeutlichung der Anteil der verwertbaren Antworten je SNR aufgetragen. Bei negativen SNRs hat die Messung der Antwortzeiten aufgrund der geringeren Datenpunkte eine geringere Aussagekraft.

Tabelle 1: Anteil der tatsächlich in das Gesamtergebnis der Antwortzeiten fließenden Antworten von insgesamt 2400 Darbietungen pro SNR. Das sind jeweils die Antworten, welche nicht mit *keine Angabe* bewertet wurden.

SNR in dB	-10	-5	0	5	10	40
Anteil verwertbarer Antworten	37.75 %	72.37 %	93.95 %	97.66 %	99.62 %	100 %

Die Betrachtung der subjektiven Bewertung der Höranstrengung anhand der 14-stufigen Bewertungsskala lieferte den in Abb. 3 dargestellten Zusammenhang. Die Werte der Gruppen, welche von den Probanden mit der 14. Stufe *nur Störgeräusch* bewertet wurden, werden hier ausgeschlossen. Dies ist bei 13 von insgesamt 120 Gruppen bei -10 dB SNR der Fall gewesen. Die den Boxplots zu Grunde liegenden Daten setzen sich aus den über je drei Listen pro SNR je Proband gemittelten Bewertungen zusammen. Über die Daten von -10 bis 10 dB SNR wird die rote Ausgleichsgerade berechnet, welche die Abszisse bei 14.5 dB SNR schneidet. Demnach werden SNRs darüber hinaus von den Probanden als mühelos empfunden.

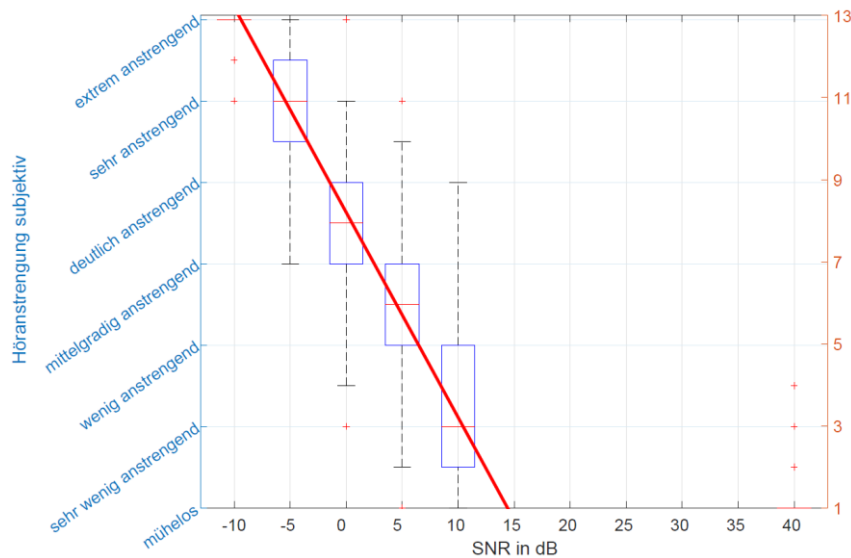


Abbildung 3: Ergebnisse der subjektiven Bewertungsskala aller Probanden als Boxplots über SNR. Ausreißer sind mit ‚+‘ gekennzeichnet. In rot ist die Ausgleichsgerade durch die Medianwerte dargestellt.

Die Ergebnisse der drei verschiedenen Erhebungen sind in Abb. 4 zusammengefasst. Die psychometrische Funktion des Sprachverstehens wurde hier mittels Spline approximiert (blaue Kurve) für eine gleichgewichtete Gegenüberstellung mit den Antwortzeiten. Die 13 auswertbaren Stufen der Bewertungsskala sind hier auf 0- 100 % skaliert, wobei 0 % der Bewertung *mühelos* und 100 % der Bewertung *extrem anstrengend* entsprechen. Auffallend ist, dass die subjektive Bewertung im Bereich zwischen 0 dB und 10 dB die Anstrengung als höher bewertet als es die Antwortzeiten indizieren. Oberhalb des SNRs mühelosen Sprachverstehens ab 14,5 dB sind jedoch noch Änderungen der Antwortzeiten zu verzeichnen. Die Korrelation zwischen subjektiver Bewertung und Antwortzeiten beträgt 71 %***. Die Korrelation zwischen subjektiver Bewertung und Sprachverstehen beträgt - 90 %***. Der Wert legt nahe, dass die subjektive Bewertung kaum mehr Informationen beinhaltet als das Sprachverstehen.

Der Vergleich der Funktionen von Antwortzeiten und Sprachverstehen zeigt ein antiproportionales Verhalten. Bei höherem SNR ist eine frühere Sättigung des Sprachverstehens gegenüber den Antwortzeiten zu verzeichnen. Diese Tatsache und die Korrelation zwischen beiden Größen von -75 %*** bestätigt tendenziell die Vermutung, dass die Antwortzeiten insbesondere im Sättigungsbereich des Sprachverstehens zusätzliche Informationen beinhalten. Es ist weiterhin zu bedenken, dass sich die Mittelungen der Antwortzeiten bei negativen SNRs aus weniger Daten zusammensetzen.

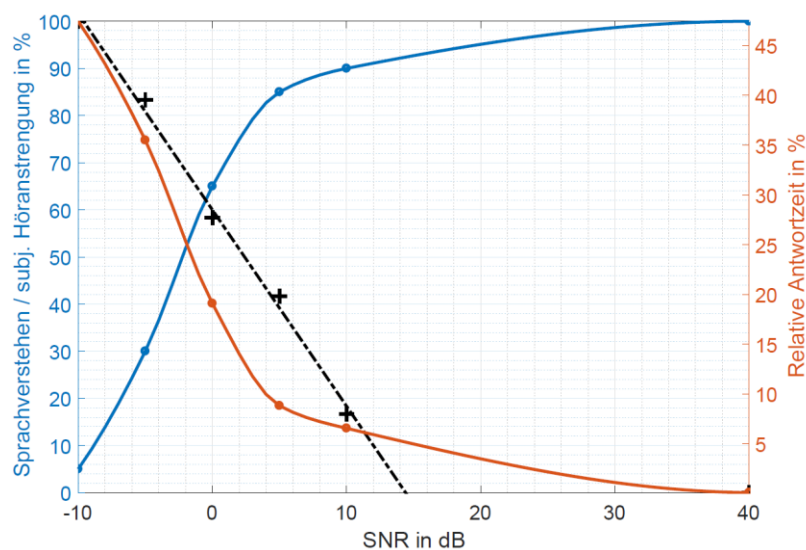


Abbildung 4: Gegenüberstellung der einzelnen Messungen. Die blaue Kurve zeigt die mittels Spline approximierten psychometrische Funktion des Sprachverstehens über SNR (Ausrichtung linke Ordinate). Die schwarz gestrichelte Linie stellt

die Ausgleichsgerade der subjektiven Höranstrengung über SNR in % dar (Ausrichtung linke Ordinate). Die Mediane der Höranstrengung sind als ‚+‘ gekennzeichnet. An der rechten Ordinate ausgerichtet ist die Spline- Approximation der relativen Antwortzeiten über SNR in Rot.

Fazit und Ausblick

Mehrfach wurde bereits ein Zusammenhang zwischen den Antwortzeiten bei Sprachtests und der Höranstrengung dokumentiert. Diese Untersuchung unterlag der Fragestellung, ob sich dieser Zusammenhang auch mit dem FET zeigen lässt. Dabei wurde Wert darauf gelegt, eine leicht in die alltägliche Praxis zu überführende Messmethode zu untersuchen. Die Messung bestätigte, dass sich durch die Erhöhung des SNRs beim FET kürzere Antwortzeiten ergeben. Die rein Software- basierte Erfassung der Antwortzeiten mittels Zeitdifferenz aus Drücken („Click“) auf die Darbietung und Drücken („Click“) auf die Antwort zeigt eine hohe Korrelation mit der automatischen Erfassung der Antwortzeit mithilfe der Mikrofonaufnahme. Das Verfahren ließe sich damit ohne viel Aufwand als Softwareupdate in bestehenden Audiometern nachrüsten. Hierbei ist es auch wichtig anzumerken, dass dadurch keine geänderte Einweisung der zu untersuchenden Person erfolgen muss. Allerdings eignet sich die Messung schlecht für negative SNRs, da sich die Menge der verwertbaren Daten hier deutlich verringert.

Weiterhin ist zu untersuchen, ob sich die Messung der Antwortzeit eignet, um etwa Vergleiche zwischen verschiedenen Hörgerätemerkmalen zu untersuchen. Insbesondere solcher Merkmale, deren Nutzen sich mutmaßlich positiv auf die Höranstrengung, jedoch nicht auf das Sprachverstehen auswirken. Derzeit wird eine Studie durchgeführt, bei welcher der Nutzen einer Störgeräuschunterdrückung bei Hörgeräteträgern mit N3-Hörverlust mit dieser Messung überprüft werden soll.

In Fällen, in denen das Sprachverstehen keine Rolle spielt, wäre es außerdem interessant zu untersuchen, ob sich die Aussagekraft der Messung der Antwortzeiten durch eine Erhöhung der kognitiven Last erhöhen könnte. Dies könnte beispielsweise durch eine veränderte Einweisung geschehen, bei der die Probanden gebeten werden, so schnell wie möglich zu antworten.

Quellen

- Hyönä J, Tammola J, Alaja A, „Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks“, *The Quarterly Journal of Experimental Psychology Section A*, vol. 48, pp. 598-612, 1995.
- Bernarding C, Strauss DJ, Hannemann R, Seidler H, Corona-Strauss FI, „Neurodynamic evaluation of hearing aid features using EEG correlates of listening effort“, *Cognitive Neurodynamics*, vol. 1, pp. 203-215, 2017.
- Sarapalis A, Kalluri S, Edwards B, Hafter E, „Objective measures of listening effort: Effects of background noise and noise reduction“, *Journal of Speech, Language and Hearing Research*, vol. 52, pp. 1230-1240, 2009.
- Krüger M, „Entwicklung und Evaluation einer adaptiven Skalierungsmethode zur Ermittlung der subjektiven Höranstrengung“, Masterarbeit, Universität Oldenburg, 2015.
- Gatehouse S, Gordon J, „Response times to speech stimuli as measures of benefit from amplification“, *British Journal of Audiology*, vol. 24, pp. 63-68, 1990.