

Der Einfluss des Präsentations- und Antwortformates auf die Messung der Sprachverständlichkeit mit Einsilbertestverfahren

Helga Sukowski¹, Thomas Brand¹, Kirsten C. Wagener², Birger Kollmeier^{1,2}

¹Medizinische Physik, Universität Oldenburg, 26111 Oldenburg

²Hörzentrum Oldenburg, Marie-Curie-Str. 2, 26129 Oldenburg

Einleitung

Im Rahmen der Begutachtung von Hörstörungen wird zur Ermittlung des prozentualen Hörverlustes aus dem Sprachaudiogramm seit langer Zeit der Freiburger Sprachtest (Hahlbrock, 1953) eingesetzt. Dieser Test besteht aus einem Zahlentest und einem Einsilbertest. In der Begutachtung wird mit dem Zahlentest die 50%-Sprachverständlichkeitsschwelle (L50: Pegel, bei dem 50% der Wörter verstanden werden) ermittelt und daraus der „Hörverlust für Zahlen“ berechnet. Mit dem Einsilbertest werden in der Begutachtung Verständlichkeiten bei 60, 80 und 100 dB SPL gemessen. Alle Messungen werden in Ruhe durchgeführt.

Der Einsilbertest des Freiburger Sprachtests steht vor allem wegen der nachgewiesenen ungleichen Schwierigkeit der Testlisten (Gruppen) seit langem in der Kritik (vgl. z.B. Bangert, 1980; von Wedel, 1986). Mit den verkürzten Testlisten des Einsilber-Reimtests nach von Wallenberg und Kollmeier (1989, „WaKo-Test“; Brand & Wagener, 2005) steht ein alternatives, gut evaluiertes Verfahren mit gleichschwierigen Testlisten zur Verfügung. In bisherigen Untersuchungen mit schwerhörenden Probanden wurde der Freiburger Einsilbertest mit dem WaKo Einsilber-Reimtest verglichen (Sukowski et al., 2007; Sukowski et al., 2008). Dabei hat sich gezeigt, dass bei gleichen Präsentationspegeln beim Einsilber-Reimtest signifikant höhere Verständlichkeiten erreicht werden als beim Freiburger Einsilbertest. Dieser Unterschied besteht auch, wenn die Reimtestergebnisse um die Ratewahrscheinlichkeit (resultierend aus dem geschlossenen Antwortformat des Tests: die Antwort wird jeweils aus fünf Alternativen ausgewählt) korrigiert sind. Wird der WaKo-Test bei einem um 15 dB reduzierten Pegel eingesetzt, zeigen sich ähnliche Ergebnisse wie mit dem Freiburger Einsilbertest. Damit wurde nachgewiesen, dass der WaKo-Test als moderneres Verfahren den Freiburger Einsilbertest grundsätzlich ersetzen kann. Bei der Anwendung in der Begutachtung ohne Veränderung der Bewertungstabellen ist unbedingt die Pegelreduktion zu beachten, da ansonsten die Hörbeeinträchtigungen der Probanden/Patienten erheblich unterschätzt würden.

In der aktuellen Studie wurde der Frage nachgegangen, wie die Unterschiede zwischen den Sprachverständlichkeitsergebnissen beider Verfahren erklärt werden können. Da sich die beiden Tests im Präsentations- und Antwortformat unterscheiden, wurde zunächst der Einfluss dieser Faktoren auf die Sprachverständlichkeit untersucht. Der Freiburger Einsilbertest ist ein offenes Verfahren, bei dem die Probanden das präsentierte Wort als Antwort wiederholen. Die Testwörter werden ohne einen Ankündigungssatz dargeboten. Der WaKo-Test ist ein geschlossenes Testverfahren, bei dem zu jedem Testwort fünf Antwortalternativen, die sich jeweils nur in einem Phonem unterscheiden, auf einem Bildschirm präsentiert werden. Außerdem geht jedem Testwort ein Ankündigungssatz voraus (z.B.: „Bitte markieren Sie das Wort TAG“). Die konkrete Frage in dieser Studie war: Welchen Einfluss haben die beiden Faktoren „Ankündigung“ und „Auswahl“ auf die gemessene Sprachverständlichkeit?

Die Hypothesen sind: (1) Werden die Testwörter mit Ankündigungssatz präsentiert, werden höhere Verständlichkeiten ermittelt als wenn die Testwörter ohne Ankündigungssatz dargeboten werden. Es wird angenommen, dass der Ankündigungssatz die Aufmerksamkeit für die Darbietung des Testwortes erhöht und es somit besser erkannt werden kann. (2) Wird das Testwort aus mehreren Alternativen ausgewählt, führt dies zu höheren Verständlichkeiten als wenn das Testwort ohne Auswahlmöglichkeit im offenen Verfahren wiederholt wird. Aufgrund der sichtbaren Alternativen kann nach einem Ausschlussprinzip vorgegangen werden und ggf. kann bereits durch die korrekte Identifikation eines einzelnen Phonems das Testwort erkannt werden.

Methode

Um den Einfluss der Faktoren „Ankündigung“ und „Auswahl“ untersuchen zu können, wurden die Originaltestversionen beider Verfahren hinsichtlich dieser Faktoren modifiziert. Insgesamt wurden die folgenden sechs Einsilbertest-Varianten konstruiert und eingesetzt. Der Ausdruck in den runden Klammern ist jeweils die Abkürzung, die in Abbildung 1 für die Varianten verwendet wird:

- Freiburger Einsilbertest ohne Ankündigungssatz (F ohne Ank) [Originalversion]
- Freiburger Einsilbertest mit Ankündigungssatz (F mit Ank)
- Einsilber-Reimtest ohne Ankündigung ohne Auswahl (R ohne Ank ohne Ausw)
- Einsilber-Reimtest mit Ankündigung ohne Auswahl (R mit Ank ohne Ausw)
- Einsilber-Reimtest ohne Ankündigung mit Auswahl (R ohne Ank mit Ausw)
- Einsilber-Reimtest mit Ankündigung mit Auswahl (R mit Ank mit Ausw) [Originalversion]

Die Untersuchung wurde in zwei Teilstudien mit unterschiedlichen Stichproben (Teil 1: Normalhörende (NH), Teil 2: geringgradig Schwerhörige (SH)) durchgeführt. In beiden Teilstudien wurden alle Probanden mit allen sechs Varianten bei jeweils fünf verschiedenen Pegeln getestet (30 Messungen pro Person, aufgeteilt auf zwei Termine). Das Sprachmaterial wurde jeweils in Ruhe dargeboten und freifeldentzerrt über Kopfhörer (Sennheiser HDA 200) präsentiert.

In Teilstudie 1 nahmen 12 normalhörende Probanden (9 Frauen, 3 Männer; Alter 21-43 Jahre) teil. Die Darbietungspegel für jede der sechs Testvarianten waren 13, 21, 29, 37, 45 dB SPL. In Teilstudie 2 bestand die Stichprobe aus 16 Personen (5 Frauen, 11 Männer; Alter: 42-66 Jahre) mit geringgradigen Hörverlusten. Die Hörverluste hatten ein Muster, das dem einer (beginnenden) Lärmschwerhörigkeit ähnlich ist (Hochtonverlust: Hörverlust bei 1000 Hz \leq 20 dB, Hörverlust bei 4000 Hz \geq 40 dB, ggf. eine Senke bei 4000 Hz). In dieser Gruppe wurden für alle Freiburger Testvarianten die Darbietungspegel 29, 37, 45, 53 und 61 dB SPL und für alle Reimtestvarianten die Darbietungspegel 21, 29, 37, 45 und 53 dB SPL verwendet.

Ergebnisse

Der Einfluss der Faktoren „Ankündigung“ und „Auswahl“ wurde mit Varianzanalysen mit Messwiederholung und den zugehörigen Posthoc-Tests berechnet. Außerdem wurden *t*-Tests für gepaarte Stichproben durchgeführt. In die Varianzanalysen wurden entweder alle sechs Testvarianten einbezogen oder nur die vier Reimtestvarianten. In der Gruppe der geringgradig Schwerhörigen gingen in die Analysen, die alle sechs Testvarianten berücksichtigten, aufgrund der unterschiedlichen Messpegelbereiche nur die Ergebnisse bei den Darbietungspegeln 29, 37, 45, 53 dB SPL ein.

Die Varianzanalysen zeigten in beiden Teilstichproben sehr ähnliche Ergebnisse. In den Berechnungen mit allen sechs Testvarianten wurde in beiden Teilstudien kein signifikanter Haupteffekt des Ankündigungssatzes ($p > 0,05$) ermittelt. Die Analysen, in denen nur die vier Reimtestvarianten (Ergebnisse jeweils um die Ratewahrscheinlichkeit korrigiert) berücksichtigt wurden, zeigten einen signifikanten Haupteffekt der Auswahlmöglichkeit ($p < 0,01$). In beiden Teilstichproben wurden in den Varianten mit Auswahlmöglichkeit signifikant höhere Verständlichkeiten ermittelt als in den Varianten ohne Auswahlmöglichkeit (NH: $M_{\text{ohneAusw}} = 69\%$, $M_{\text{mitAusw}} = 83\%$; SH: $M_{\text{ohneAusw}} = 50\%$, $M_{\text{mitAusw}} = 71\%$). In diesen Analysen wurde zwar in beiden Gruppen auch ein signifikanter Effekt für den Faktor „Ankündigung“ nachgewiesen. Die Unterschiede der Mittelwerte waren jedoch sehr gering (NH: $M_{\text{ohneAnk}} = 75\%$, $M_{\text{mitAnk}} = 77\%$; SH: $M_{\text{ohneAnk}} = 60\%$, $M_{\text{mitAnk}} = 62\%$).

Insgesamt wurde mit den Ergebnissen auch die bisherige Beobachtung bestätigt, dass bei gleichen Darbietungspegeln mit dem Einsilber-Reimtest (alle Varianten) signifikant höhere Verständlichkeiten ermittelt werden als mit dem Freiburger Einsilbertest (alle Varianten).

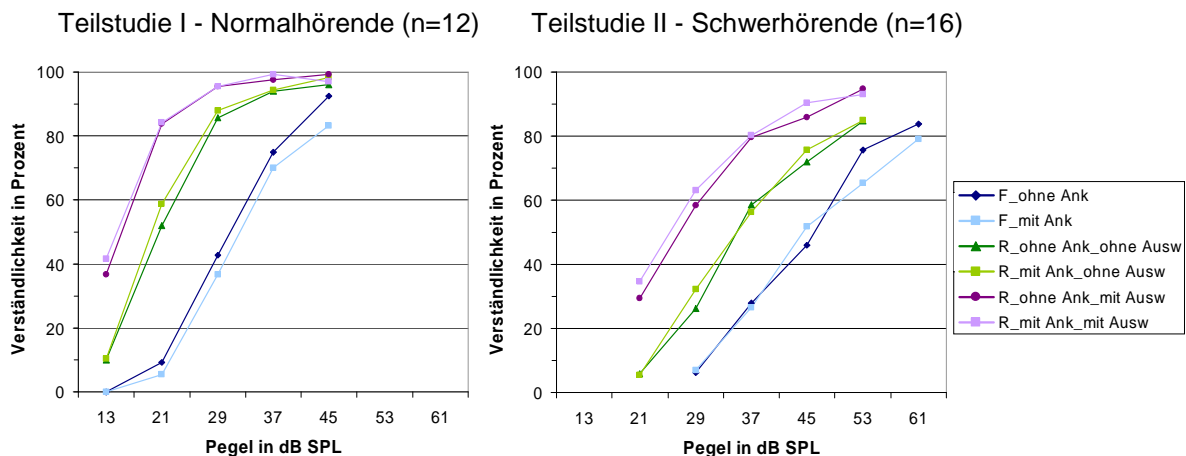


Abb. 1: Mittelwerte für alle sechs Testvarianten bei jeweils fünf Darbietungspegeln, links für die Gruppe der normalhörenden Probanden und rechts für die Gruppe der Probanden mit geringgradigen Hörverlusten.

Um Vergleiche zwischen verschiedenen Testvarianten auch über beide Teilstudien hinweg vornehmen zu können, wurden durch Anpassung einer logistischen Diskriminationsfunktion (Brand & Kollmeier, 2002) für alle Probanden die 50%-Sprachverständlichkeitsschwelle sowie die Steigung an diesem Punkt berechnet. Auf der Grundlage der berechneten Sprachverständlichkeitsschwellen wurden gezielt Vergleiche ausgewählter Testvarianten durchgeführt. Zwei Vergleiche waren von besonderer Bedeutung:

(1) *Freiburger Einsilbertest ohne Ankündigung - Einsilber-Reimtest ohne Ankündigung ohne Auswahl*: Dieser Vergleich ist wichtig, da mögliche Unterschiede nicht auf die beiden untersuchten Faktoren (Auswahl und Ankündigung) zurückgeführt werden können. *t*-Tests für abhängige Stichproben, die getrennt für beide Teilstudien durchgeführt wurden, zeigten erwartungsgemäß in beiden Probandengruppen einen signifikanten Unterschied zwischen den ausgewählten Testvarianten. Die Differenzen der Mittelwerte betragen in der Gruppe

NH 10 dB und in der Gruppe SH 8,8 dB. Der Unterschied zwischen beiden Gruppen war nicht signifikant ($p=0,27$).

(2) *Einsilber-Reimtest ohne Auswahl - Einsilber-Reimtest mit Auswahl (jeweils mit Ankündigung)*: Bei diesem Vergleich ist die Frage von Bedeutung, ob sich ein einheitlicher Wert bestimmen lässt, um den der Pegel bei der offenen Anwendung angehoben werden sollte, damit die gleichen Ergebnisse wie bei der geschlossenen Anwendung erzielt werden. Bezüglich dieser Frage wurden ebenfalls getrennt für beide Teilstudien *t*-Tests für abhängige Stichproben zum Vergleich der 50%-Sprachverständlichkeitsschwellen beider Testvarianten durchgeführt. In beiden Teilstudien wurde wiederum erwartungsgemäß ein signifikanter Unterschied zwischen den ausgewählten Testvarianten ermittelt (Differenzen der Mittelwerte: NH: 7,5 dB; SH: 11,6 dB). Der Unterschied zwischen beiden Probandengruppen war bei diesem Vergleich signifikant ($p=0,02$).

Diskussion

Die Berechnungen mit allen sechs Testvarianten haben keinen signifikanten Effekt für den Ankündigungssatz gezeigt. Das heißt, in den beiden hier untersuchten Probandengruppen wurden entgegen der Erwartung keine höheren Sprachverständlichkeiten ermittelt, wenn die Wörter mit einem Ankündigungssatz präsentiert wurden. Obwohl ein signifikanter Haupteffekt ausgeblieben ist, sollte der Einfluss der Ankündigung auf die gemessene Verständlichkeit nicht pauschal zurückgewiesen werden. Die Berechnungen, die lediglich die vier Reimtestergebnisse einschlossen, hatten zwar in absoluten Zahlen geringe, aber dennoch signifikante Mittelwertsunterschiede zwischen den Varianten mit und ohne Ankündigung gezeigt. Außerdem ist zu vermuten, dass Personen mit stärkeren Hörverlusten deutlicher von einem Ankündigungssatz profitieren können als dies in den hier untersuchten Gruppen der Fall war. Weitere Untersuchungen zu dieser Frage sind daher geplant.

Die Auswahlmöglichkeit in den Reimtestvarianten hat sich als relevanter Faktor herausgestellt. In den Varianten, in denen die Auswahlmöglichkeit gegeben war, wurden unabhängig vom Ankündigungssatz jeweils höhere Verständlichkeiten ermittelt als in den Varianten, in denen das Testwort wie beim Freiburger Test nachgesprochen werden sollte. Das bedeutet, Auswahlalternativen am Bildschirm zu sehen erhöht die Trefferquote für das Zielwort und somit schließlich die gemessene Sprachverständlichkeit.

Bei den Berechnungen, die auf der Basis der 50%-Sprachverständlichkeitsschwellen durchgeführt wurden, hatte sich angedeutet, dass die schwerhörenden Probanden etwas stärker von der Auswahlmöglichkeit profitieren konnten als die normalhörenden Probanden. Es ist nicht auszuschließen, dass dieser Vorteil bei höhergradig Schwerhörenden noch zunimmt. Auf der Basis der bislang untersuchten Stichproben lässt sich daher noch kein einheitlicher Wert festlegen, um den der Pegel bei einer offenen Anwendung des Reimtests anzuheben wäre, damit ähnliche Ergebnisse wie bei der geschlossenen Anwendung erzielt werden. Zu diesem Punkt wären ebenfalls weitere Studien mit den entsprechenden Stichproben notwendig.

Die vorliegende Studie hat gezeigt, dass die beiden untersuchten Faktoren „Ankündigung“ und „Auswahl“ einen Teil der Differenz zwischen den beiden Originaltestversionen erklären können. Sowohl bei Normalhörenden als auch bei Schwerhörenden bleibt jedoch bei gleichem Präsentations- und Antwortformat ein Restunterschied in den Sprachverständlichkeitsergebnissen bestehen (ca. 9-10 dB von ursprünglich 18 dB (NH) bzw. 21 dB (SH) bezogen auf den L50). Zur Erklärung der Unterschiede zwischen dem Einsilber-Reimtest und dem Freiburger Einsilbertest müssen daher noch andere Faktoren wie beispielsweise die Aufnahmequalität, die Definition des Sprachpegels oder Sprechereffekte diskutiert werden. Der Einfluss dieser Faktoren lässt sich allerdings nicht durch einfache Modifikationen bei der Durchführung überprüfen. Da sich der Einfluss weiterer Faktoren jedoch offenbar in beiden Teilstichproben in gleicher Weise auswirkt, stellen die daraus resultierenden Verständlichkeitsunterschiede kein Problem für weitere Vergleichsmessungen mit den beiden Einsilbertestverfahren dar.

Unterstützt durch das MWK „Audiologie-Initiative Niedersachsen“; V2.22 04032 (Audiologische Forschung).

Literatur

- Bangert H (1980) Probleme bei der Ermittlung des Diskriminationsverlustes nach dem Freiburger Sprachtest. *Audiol Akustik* 19: 166–170
- Brand T, Kollmeier B (2002) Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. *J Acoust Soc Am* 111: 2801–2810
- Brand T, Wagener K (2005) Wie lässt sich die maximal erreichbare Verständlichkeit optimal bestimmen? In: 8. Jahrestagung der Deutschen Gesellschaft für Audiologie, Göttingen, 24.–26.2.2005, CD-ROM
- Hahlbrock KH (1953) Über Sprachaudiometrie und neue Wörterteste. *Arch Ohren Nasen Kehlkopfheilkd* 162: 394–431
- Sukowski H, Brand T, Wagener KC & Kollmeier B (2008) Untersuchung zur Vergleichbarkeit des Freiburger Sprachtests mit dem Göttinger Satztest und dem Einsilber-Reimtest nach von Wallenberg und Kollmeier. HNO, im Druck

- Sukowski H, Brand T, Wagener KC & Kollmeier B (2007) Sprachverständlichkeitstests in Ruhe: Gibt es alternative Verfahren zum Freiburger Sprachtest in der Begutachtung bei (Lärm-) Schwerhörigkeit? In: Fortschritte der Akustik - DAGA 2007, DEGA e.V., Berlin, S. 715-716
- Wallenberg E-L von, Kollmeier B (1989) Sprachverständlichkeitsmessungen für die Audiologie mit einem Reimtest in deutscher Sprache: Erstellung und Evaluation von Testlisten. *Audiol Akustik* 28: 50–65
- Wedel H von (1986) Untersuchungen zum Freiburger Sprachtest - Vergleichbarkeit der Gruppen im Hinblick auf Diagnose und Rehabilitation (Hörgeräteanpassung und Hörtraining). *Audiol Akustik* 25: 60–73