



Die adaptive Messung des Freiburger Einsilbertests im Störschall

Entwicklung einer Messmethode und Vergleich der Ergebnisse mit dem Oldenburger Satztest

Hintergrund

Für die Ermittlung des Schweregrads einer Schwerhörigkeit werden gemeinhin ton- und sprachaudiometrische Messungen durchgeführt sowie Fragebögen wie z. B. der APHAB-Fragebogen (Abbreviated Profile of Hearing Aid Benefit) verwendet, welche die subjektive Beeinträchtigung des Hörvermögens erfassen [11].

In der Sprachaudiometrie wird das Sprachverstehen ermittelt, indem die präsentierten Wörter nachgesprochen werden müssen und vom Untersucher auf Gleichheit bewertet werden. Darin unterscheidet sich ein Sprachverstehenstest von einem Sprachsinverständnis, bei dem das Begreifen eines sprachlichen Inhalts gemessen wird, von einem Sprach(entwicklungs)test bei einem Kind sowie von einem Sprachverständlichkeitstest, bei dem die Verständlichkeit der Aussprache eines Patienten beurteilt wird (wobei der Untersucher der Hörer ist, im Gegensatz zur Audiometrie, bei welcher der Patient der Hörer ist). Daher wird im Folgenden der Begriff Sprachverstehenstest verwendet.

Zur Überprüfung des Sprachverstehens existiert eine Vielzahl von Tests, die sich anhand ihres Aufbaus in Wort- und

Satztests unterscheiden lassen. Weitere Unterscheidungsmerkmale sind eine adaptive oder nichtadaptive Pegelanpassung und die Anwendung der Sprachverstehenstests ohne und mit Störschall. Im deutschsprachigen Raum etablierte sich unter den Worttests seit seiner Veröffentlichung im Jahr 1953 der Freiburger Sprachtest [23]. Dieser setzt sich aus einem Einsilber- und einem Zahlentest zusammen und verwendet ein nicht-adaptives Messverfahren [13]. Aufgrund seiner einfachen und schnellen Durchführbarkeit ist er der meistverwendete Sprachverstehenstest im deutschsprachigen Raum.

Von den verschiedenen Satztests werden hauptsächlich der jeweils adaptiv gemessene Göttinger Satztest (GÖSA) und der Oldenburger Satztest (OLSA) verwendet [11]. Hierbei wird der Sprachschallpegel ermittelt, bei dem während 65 dB Störschallpegel das Sprachsignal zu 50 % verstanden wird. Die Differenz zwischen dem Sprachschallpegel für das 50 %-Sprachverstehen und dem Störschallpegel ist als Schwelle L50 für das Sprachverstehen definiert und wird im Folgenden als Signal-Rausch-Abstand für das 50 %-Sprachverstehen bezeichnet (S/N50). Im Vergleich mit dem Freiburger Einsilbertest (FBE) oder dem

GÖSA ist die Durchführung des OLSA zeitaufwendiger, da zur Vermeidung von Trainingseffekten zunächst Trainingsmessungen durchgeführt werden müssen.

Jedoch weist der OLSA eine im Vergleich mit dem FBE geringere Streuung der Ergebnisse auf. Dies ist durch die für den jeweiligen Sprachverstehenstest spezifische Diskriminationsfunktion zu erklären. Der Zusammenhang zwischen Sprachschallpegel und Sprachverstehen wird durch die sigmoidale Diskriminationsfunktion beschrieben [15]. Der OLSA misst das 50 %-Sprachverstehen am Punkt der größten Steigung, dem Wendepunkt der Diskriminationsfunktion. Mit dem FBE wird das bei einem bestimmten Sprachschallpegel erhaltene maximale Einsilberverstehen bestimmt. Im Verlauf der Diskriminationsfunktion liegt das maximale Einsilberverstehen im dem Wendepunkt folgenden Sättigungsbereich der Diskriminationsfunktion. Durch die Redundanz der Testwörter des OLSA ist die Steigung der Diskriminationsfunktion im Wendepunkt größer, wodurch der OLSA im Vergleich eine geringere Streuung der Ergebnisse aufweist. Zudem wirkt sich Ratewahrscheinlichkeit des Basiswortmaterials und der prozedurale Lerneffekt

des OLSA auf die Verständlichkeit aus, sodass es zu einer Verschiebung der Diskriminationsfunktion und damit zu kleineren S/N50 kommt.

Allerdings verlieren Patienten erst nach mehreren Testdurchläufen des OLSA ihre Lerneffekte [28]. Erfahrene Patienten, speziell Probanden von Studien, liefern bessere Ergebnisse, die in anderen Normwerten berücksichtigt werden müssten. Zudem hängen die Ergebnisse des OLSA auch vom auditiven Arbeitsgedächtnis ab, das bei Patienten mit auditiven Verarbeitungsstörungen, unter denen auch Erwachsene und nicht nur Kinder leiden, signifikant eingeschränkt ist [28]. Diese Lerneffekte wirken sich auch nach vielen Wochen noch aus [28].

Beide Sprachverstehenstests werden auch im Störschall angewendet [24, 25, 34]. Die Störschallmessung ist wegen der Forderung nach maximalem Behinderungsausgleich im Rahmen einer Hörgeräteversorgung nach den „tragenden Gründen“ der Hilfsmittel-Richtlinie von besonderer Bedeutung [12]. Die Anwendbarkeit des FBE im Störschall (FBE-S) war Gegenstand einer kontroversen Diskussion [19, 21, 32]. In der jüngeren Vergangenheit wurden verschiedene Untersuchungen zur Anwendung im Störschall durchgeführt [10, 24–26]; nach der Evaluierung [24] erfolgte im Jahr 2017 die Aufnahme des FBE-S in die Hilfsmittel-Richtlinie [11]. Zur Optimierung der Messgenauigkeit empfiehlt die Hilfsmittel-Richtlinie beim FBE-S die Verwendung von 2 Listen [11].

Nach Einführung des Freiburger Sprachtests wurden insbesondere bezüglich des FBE verschiedene Studien zu dessen Validität, Reliabilität und Effizienz veröffentlicht [1]. In der jüngsten Diskussion wurde u. a. auf die phonemische Ausgewogenheit [9], die Test-Retest-Reliabilität [35] oder einen möglichen Trainingseffekt des FBE eingegangen [29]. Die Kritik am FBE führte zu Vergleichen mit alternativen Sprachverstehenstests. So wurde der FBE als Worttest mit seiner bisher prozentualen Messweise mit adaptiv messenden Satztests wie dem GÖSA oder dem OLSA in Ruhe oder im Störschall verglichen und aus vergleichenden Messungen eine

Testäquivalenz abgeleitet [20, 31]. Die direkte Vergleichbarkeit verschiedener Sprachverstehenstests ist jedoch besonders bei unterschiedlichen Messweisen nur eingeschränkt möglich [15].

In der bisherigen Diskussion wurde die Möglichkeit einer adaptiven Pegelsteuerung des FBE nicht untersucht. Gegen eine adaptive Messung des FBE (aFBE), bei der analog zum OLSA oder verwandten Tests ebenfalls der Signal-Rausch-Abstand für das 50 %-Sprachverstehen berechnet würde, sprach argumentativ im Wesentlichen die Unausgewogenheit der Testlisten [14]. In der Diskussion um einen prozeduralen Trainingseffekt des FBE wurde als Begründung für den Anstieg des mittleren Sprachverstehens zwischen 2 Messdurchläufen die Habituation an die Testsituation angeführt [28]. Bisher konnte keine Aussage darüber getroffen werden, ob der Trainingseffekt durch Habituation auch bei vorheriger Anwendung eines anderen Sprachverstehenstests wie dem OLSA eintritt und sich somit die Präsentationsreihenfolge beim Vergleich zweier Tests auf die Ergebnisse auswirken kann.

Angesichts der bedingten Vergleichbarkeit von Sprachverstehenstests bei unterschiedlichen Messweisen soll mit dieser Studie ein Vorschlag zu einer adaptiven Messweise des FBE im Störschall vorgestellt werden (aFBE-S). Die Ergebnisse des aFBE-S werden mit denen des adaptiv messenden OLSA im Störschall (OLSA-S) verglichen und hinsichtlich der Messgenauigkeit, des Einflusses der Testreihenfolge und aufgrund fehlender Werte in der Literatur bezüglich des Zeitbedarfs untersucht. Des Weiteren sollen mögliche geschlechtsspezifische Unterschiede der Ergebnisse untersucht werden. Verschiedene Analysen haben gezeigt, dass sich die meisten kognitive Funktionen zwischen Frauen und Männern nicht unterscheiden [18], jedoch ist ein geschlechtsspezifischer Unterschied der Ergebnisse in diesem Vergleich nicht auszuschließen. Zudem sollen die S/N50 der verwendeten Sprachverstehenstests auf eine Korrelation überprüft werden.

Diese Untersuchung ist Teil des Evidenzprojekts des Deutschen Studienzentrums HNO.

Methoden

An der Studie beteiligten sich 40 Probanden (25 w., 15 m.) im Alter zwischen 18 und 25 Jahren (Mittelwert: $22,0 \pm 1,9$ Jahre). Neben dem Alter galten als Einschlusskriterien Deutsch als Muttersprache und ein beidseitiges normales Hörvermögen. Alle Studienteilnehmer erteilten schriftlich ihr Einverständnis zur Teilnahme. Die Messungen erfolgten in der Sektion für Phoniatrie und Pädaudiologie des Universitätsklinikums Schleswig-Holstein, Campus Lübeck, in einem nach DIN-Norm schallarmen Raum unter Verwendung eines nach DIN-Norm kalibrierten klinischen Audiometers AT900 (Fa. Auritec Medizindiagnostische GmbH, Hamburg, Deutschland) [5–8]. Um den Einfluss verschiedener Instruktionen auf die Messungen zu verringern, wurden einheitliche Erklärungen zum aFBE-S gegeben. Diese orientierten sich an den Instruktionen, welche im Handbuch des OLSA auf einer Audio-CD beschrieben sind [16]. Die Reihenfolge der Messung mit dem aFBE-S und dem OLSA-S erfolgte randomisiert. So wurden die Messungen 20-mal mit dem OLSA und 20-mal mit dem aFBE-S begonnen. Die Messungen wurden als binaurale Freifeldmessungen durchgeführt. Die Lautsprecher befanden sich frontal im Abstand von 1 m zur Kopfmitte (0° -Richtung), Sprache und Störschall wurden beim aFBE-S und dem OLSA-S aus demselben Lautsprecher präsentiert (S_0N_0 -Situation). Die Probanden wurden gebeten, während der Messungen eine aufrecht sitzende Position einzunehmen und sich nicht mit dem Kopf den Lautsprechern zu nähern oder diesen zur Seite zu drehen. Während der Messungen wurde die Dauer der Testdurchführung erfasst. Beim OLSA-S startete die Messung des Zeitbedarfs mit der Anwendung einer Trainingsliste, auf die eine Testliste folgte. Beim aFBE-S wurde keine Trainingsliste durchgeführt. Bei beiden Sprachverstehenstests wurde das jeweilige Störgeräusch kontinuierlich und mit einem festen Störpegel von 65 dB SPL präsentiert.

HNO 2019 · 67:118–125 <https://doi.org/10.1007/s00106-018-0597-z>
 © Der/die Autor(en) 2018

T. Memmeler · R. Schönweiler · B. Wollenberg · J. Löhler

Die adaptive Messung des Freiburger Einsilbertests im Störschall. Entwicklung einer Messmethode und Vergleich der Ergebnisse mit dem Oldenburger Satztest

Zusammenfassung

Hintergrund. Während sich die adaptive Messweise unter den Satztests etabliert hat, ist sie bei dem Freiburger Einsilbertest (FBE) als wichtigstem Vertreter der Worttests noch nicht angewendet worden. Da sich die Ergebnisse von Wort- und Satztests wegen der unterschiedlichen Messweise nicht vergleichen lassen, haben wir in dieser Studie einen adaptiven Messalgorithmus auf den FBE im Störschall übertragen, der sich bei Satztests bewährt hat. Die Ergebnisse des neuen Messverfahrens haben wir mit denen des Standards unter den Satztests, dem Oldenburger Satztest (OLSA), verglichen.

Methoden. An 40 otologisch gesunden Probanden wurde in randomisierter Reihenfolge sowohl der adaptiv gemessene FBE als

auch der OLSA durchgeführt. Die Ergebnisse wurden bezüglich ihres Zeitbedarfs, möglicher geschlechtsspezifischer Unterschiede und der Reihenfolge der Sprachverstehenstests ausgewertet und hinsichtlich der Korrelation mit dem neuen Test untersucht.

Ergebnisse. Der Zeitbedarf für den adaptiven FBE war signifikant größer als für den OLSA. Bezüglich möglicher geschlechtsspezifischer Unterschiede oder der Präsentationsreihenfolge ließen sich keine signifikanten Unterschiede der Testergebnisse nachweisen. Der mittlere Signal-Rausch-Abstand für das 50%-Sprachverstehen des OLSA war signifikant kleiner als der des FBE. Zwischen den Ergebnissen beider Tests konnte keine Korrelation nachgewiesen werden.

Schlussfolgerungen. Der FBE kann nicht nur den Diskriminationsverlust als Hörverlust in Prozent, sondern mit einem adaptiven Algorithmus auch die 50%-Sprachverstehensschwelle messen. Der adaptive FBE ist zeitaufwendiger als der OLSA. Hinsichtlich einer möglichen fehlenden Korrelation der Ergebnisse für das 50%-Sprachverstehen beider Tests sollten weitere Studien mit Schwerhörenden durchgeführt werden.

Schlüsselwörter

Sprachaudiometrie · Freiburger Einsilbertest · Oldenburger Satztest · Sprachverstehenstest · Adaptive Methode

The adaptive Freiburg monosyllabic test in noise. Development of a procedure and comparison of the results with the Oldenburg sentence test

Abstract

Background. Whereas sentence tests are commonly performed using an adaptive procedure, this method has not yet been transferred to the Freiburg monosyllabic speech test, the most important word test. When using different procedures, a comparison of results between sentence and word tests is not possible. Therefore, an adaptive procedure which has proven itself in sentence tests was transferred to the Freiburg monosyllabic test in noise. The results of the new procedure were compared to the standard of sentence tests, i.e., the Oldenburg sentence test.

Methods. The adaptive Freiburg monosyllabic speech test and the Oldenburg sentence test were applied in 40 otologically normal

subjects in a randomized order. Results were analyzed with respect to time requirements, possible gender differences, the influence of test order, and correlation of test results.

Results. The time required for the adaptive Freiburg monosyllabic speech test was significantly higher than for the Oldenburg sentence test. No significant impact of gender or test order could be shown. The mean signal-to-noise ratio for 50% speech discrimination of the Oldenburg sentence test was significantly smaller than for the adaptive Freiburg monosyllabic speech test. No correlation could be shown between the results of the two tests

Conclusion. The Freiburg monosyllabic test can not only be used for quantifying

discrimination loss in percentage terms, but also to measure the 50% speech recognition threshold with an adaptive algorithm. However, the procedure of the adaptive Freiburg monosyllabic test is more time consuming than that of the Oldenburg sentence test. Concerning a possible missing correlation between the results for 50% speech discrimination, further studies with hearing-impaired persons are needed.

Keywords

Speech audiometry · Freiburg monosyllabic speech test · Oldenburg sentence test · Speech discrimination test · Adaptive procedure

Tonaudiometrische Messung

Vor Durchführung des aFBE-S und des OLSA-S (Fa. HörTech gGmbH, Oldenburg, Niedersachsen, Deutschland) wurden bei jedem Probanden eine tonaudiometrische Messung mit einem DT 48-Kopfhörer (Fa. beyerdynamic GmbH & Co. KG, Heilbronn, Baden-Württemberg, Deutschland) durchgeführt. Der zulässige Reintonhörverlust

für die Luftleitung betrug für die Frequenzen 0,125; 0,250; 0,5; 1; 2; 3; 4; 6; 8 kHz stets maximal 10 dB HL.

Adaptive Messung des Freiburger Einsilbertests im Störschall

Für die adaptive Sprachpegeländerung wurde der Algorithmus des OLSA-S auf den FBE übertragen [16]. Das Grundprinzip dieses Algorithmus ist es, durch

Pegeländerungen den Sprachschallpegel aufzufinden, bei dem exakt 50 % der vorgespielten Wörter richtig verstanden und korrekt nachgesprochen werden. Eine Pegeländerung erfolgt abhängig von der Anzahl der richtig nachgesprochenen Wörter. Werden beispielsweise zu Beginn der Messung 4 Wörter richtig nachgesprochen, so wird der Sprachpegel im darauffolgenden Satz um 2 dB verringert, während der Störschall konstant

5er Gruppe	Anzahl der richtig verstandenen Einsilber	Pegeländerung [dB]	Sprachpegel [dB]	Testliste
1	4	-2	65	7
2	2	1	63	
3	3	-1	64	
4	1	2	63	
5	4	-1	65	
6	4	-1	64	2
7	3	0	63	
8	3	0	63	
9	3	0	63	
10	3	0	63	20
11	3	0	63	
12	5	-2	63	
13	3	0	61	
14	1	1	61	8
15	2	0	62	
16	1	1	62	
17	2	0	63	
18	2	0	63	16
19	3	0	63	
20	3	0	63	
21	2	0	63	
22	4	-1	63	9
23	3	0	62	
24	1	1	62	
25	3	0	63	
26	3	0	63	1
27	2	0	63	
28	4	-1	63	
29	3	0	62	
30	3	0	62	4
31			62	
∅			62,45	
S/N50			-2,55	
S/N50 _{OLSA-S}			-8,1	
t _{OLSA-S}			8' 30"	
t _{aFBE-S}			9' 16"	

Abb. 1 ▲ Beispielhafte adaptive Messung des Freiburger Einsilbertests im Störschall. ∅ Arithmetisches Mittel der letzten zwanzig Sprachpegel in Dezibel, S/N50 Signal-Rausch-Abstand für 50 %-Sprachverstehen in Dezibel, t_{OLSA-S}/t_{aFBE-S} benötigte Zeit in Minuten und Sekunden zur Testdurchführung im Störschall

bei 65 dB SPL bleibt. Im Verlauf wird bei 2 oder 3 richtig nachgesprochenen Wörtern der vorherige Sprachpegel belassen und keine Pegeländerung durchgeführt [16].

Die verwendeten Testlisten des FBE wurden bei jeder Messung zufällig ausgewählt, unter Ausschluss der Testlisten 5; 11; 12 und 15, welche sich als nicht perzeptiv äquivalent erwiesen [2]. Die Einsilber des FBE lagen in digitalisierter Form der Aufnahmen aus dem Jahr 1969 nach DIN 45621-1 vor [5]. Die Präsentationsreihenfolge der Einsilber blieb während

der adaptiven Messung stets unverändert. Als Störschall kam das in digitalisierter Form vorliegende sog. CCITT-Rauschen (Comité Consultatif International Téléphonique et Télégraphique) zur Anwendung [17].

Da dieser Algorithmus auf den OLSA-S ausgelegt ist, wurde der Testablauf des FBE für die adaptive Messweise angepasst. Insgesamt sind für den aFBE-S 8 Testlisten des FBE mit je 20 Einsilbern nötig. Anstatt 160 Einsilber wurden analog zum OLSA-S ebenfalls 150 Wörter dargeboten. Beim OLSA-S wird eine

Pegeländerung nach jedem Satz, somit alle 5 Wörter, durchgeführt. Aus diesem Grund wurde für die adaptive Messung des FBE das Repertoire an 150 Einsilbern in 5er-Gruppen unterteilt. Jeweils nach Präsentation von 5 Einsilbern erfolgte eine Pegeländerung gemäß Algorithmus. Im Gegensatz zum OLSA-S mussten sich beim aFBE-S die dargebotenen Wörter nicht gemerkt werden, sondern konnten unmittelbar nachgesprochen werden.

Für die Messung des aFBE-S wurde eine Microsoft-Excel-Tabelle (Fa. Microsoft Corporation, Redmond, Washington, USA) entworfen, die auf dem zuvor genannten Algorithmus basiert (Abb. 1). Zusätzlich war die Anwendungssoftware des FBE nötig, wodurch 2 PC in Gebrauch waren. Die Anzahl der richtig verstandenen Einsilber wurde während der Messung in die Excel-Tabelle eingetragen und der Sprachpegel der nächsten 5er-Gruppe in der Tabelle gemäß Algorithmus automatisch berechnet. In der Anwendungssoftware des FBE wurden die einzelnen Testlisten ausgewählt und die Pegeländerungen vorgenommen. Die Darbietung des jeweils folgenden Einsilbers erfolgte nicht in einem festen Zeitabstand, sondern erst nach Eingabe des richtig oder falsch nachgesprochenen Einsilbers durch den Versuchsleiter.

Im Verlauf der Messung werden die Sprachpegeländerungen den Ergebnissen angepasst. Bei den ersten 5 Präsentationen der 5er-Gruppen erfolgte die Pegeländerung des Sprachsignals in größeren Schritten, ab der sechsten 5er-Gruppe wurde die Pegeländerung des Sprachsignals kleiner. Die einunddreißigste 5er-Gruppe, oder Satz 31 bei der manuellen Ausführung des OLSA-S, wurde gemäß dem Algorithmus aus der Antwort der dreißigsten 5er-Gruppe berechnet. Daher wurden insgesamt 31 verschiedene Sprachpegel erfasst. Aus den Sprachpegeln der letzten zwanzig 5er-Gruppen wurde das arithmetische Mittel gebildet. Durch Bildung der Differenz zwischen dem arithmetischen Mittel der letzten 20 Sprachpegel und dem konstanten Störschallpegel von 65 dB wurde der Signal-Rausch-Abstand für das 50 %-Sprachverstehen berechnet [16].

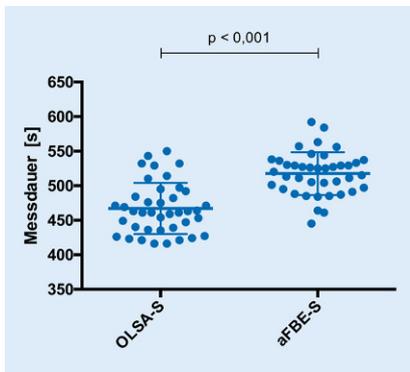


Abb. 2 ▲ Vergleich des Zeitbedarfs in Sekunden für die Messung des Oldenburger Satztests (OLSA-S) und die adaptive Messung des Freiburger Einsilbertests (aFBE-S) im Störschall

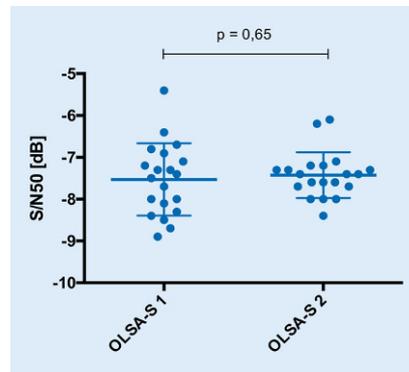


Abb. 3 ▲ Vergleich der Signal-Rausch-Abstände für das 50 %-Sprachverstehen in Dezibel (S/N50 [dB]) des Oldenburger Satztests im Störschall (OLSA-S) zwischen der Gruppe mit Präsentation des OLSA-S vor dem adaptiven Freiburger Einsilbertest (OLSA-S 1) und der Gruppe mit umgekehrter Präsentationsreihenfolge (OLSA-S 2)

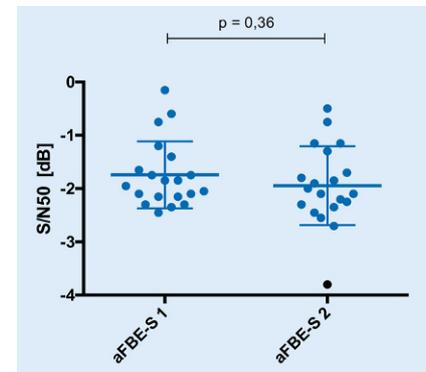


Abb. 4 ▲ Vergleich der Signal-Rausch-Abstände für das 50 %-Sprachverstehen in Dezibel (S/N50 [dB]) des adaptiven Freiburger Einsilbertests im Störschall (aFBE-S) zwischen der Gruppe mit Präsentation des aFBE-S vor dem Oldenburger Satztest (aFBE-S 1) und der Gruppe mit umgekehrter Präsentationsreihenfolge (aFBE-S 2)

Messung des Oldenburger Satztests im Störschall

Die Messung des OLSA-S (Version 1.4.6.0, Fa. HörTech gGmbH, Oldenburg, Niedersachsen, Deutschland) erfolgte mit den zu größerer Messgenauigkeit beitragenden und empfohlenen 30er-Listen, bestehend aus 30 Sätzen mit jeweils 5 Wörtern [16, 33]. Jede Messung des OLSA-S bestand aus zwei 30er-Listen im Störschall, einer Trainingsliste und der eigentlichen Testliste. Das Ergebnis der Testliste des OLSA-S wurde dokumentiert. Die Probanden sollten den vollständigen Satz oder soweit wie möglich Teile davon nachsprechen. Das Nachsprechen begann erst nach Präsentation des gesamten Satzes. Die Richtigkeit des Nachsprechens der Wörter eines Satzes wurde vom Versuchsleiter in einer Auswahlmaske der Anwendungssoftware des OLSA-S markiert. Daraufhin wurde der nächste Satz präsentiert. Die Pegeländerungen und die Berechnung des S/N50 erfolgten automatisch durch die Software des OLSA-S.

Statistische Analyse

Zur statistischen Auswertung der gewonnenen Daten wurde die Software Graph Pad Prism 6.0 (Fa. Graph Pad Software Inc., San Diego, Kalifornien, USA) verwendet. Die Ergebnisse der Messungen des aFBE-S und des OLSA-S wurden als

Streudiagramme unter Angabe der Mittelwerte mit Standardabweichung dargestellt. Die Normalverteilung der Daten wurde mit dem D'Agostino-Pearson-omnibus-K2-Test überprüft [3]. Zum Gruppenvergleich der Ergebnisse des OLSA-S und des aFBE-S abhängig von der Präsentationsreihenfolge und abhängig vom Geschlecht wurde jeweils der Welch-Test durchgeführt. Für den Vergleich des Zeitbedarfs und für den geschlechtsunabhängigen Vergleich der Signal-Rausch-Abstände für das 50 %-Sprachverstehen des OLSA-S und des aFBE-S wurde jeweils der gepaarte t-Test angewendet. Zur Untersuchung des Zusammenhangs der S/N50 wurde eine Pearson-Korrelation durchgeführt. Bei allen Gruppenunterschieden galt ein $p \leq 0,05$ als signifikant.

Ergebnisse

Zeitbedarf und geschlechtsspezifische Unterschiede

Eine Messung des OLSA-S dauerte im Durchschnitt 467 ± 37 s (7 min 47 s). Die längste Messung betrug 550 s (9 min 10 s), während die kürzeste 416 s (6 min 56 s) in Anspruch nahm (Abb. 2). Im Vergleich betrug die Dauer einer durchschnittlichen Messung des aFBE-S 518 ± 31 s (8 min 38 s). Die längste Messung des aFBE-S dauerte 592 s (9 min 52 s), die kürzeste 445 s (7 min 25 s). Hinsichtlich des Zeitbedarfs konnte ein signifikanter

Unterschied ($t = 9,29$ $p < 0,001$) nachgewiesen werden. Zwischen den untersuchten beiden Geschlechtern ließ sich kein signifikanter Unterschied bezüglich der Ergebnisse des OLSA-S und des aFBE-S nachweisen (OLSA-S: $t = 0,11$ $p = 0,91$ aFBE-S: $t = 0,89$ $p = 0,38$).

Präsentationsreihenfolge

Das Probandenkollektiv wurde in 2 Gruppen unterteilt, in denen zuerst der OLSA-S (OLSA-S 1) oder der aFBE-S (aFBE-S 1) angewendet wurden. In den Gruppen OLSA-S 2 und aFBE-S 2 kamen die Tests jeweils als Zweites zur Anwendung. Im Vergleich zeigte sich in der Gruppe OLSA-S 1 ein mittlerer S/N50 von $-7,53 \pm 0,86$ dB und in der Gruppe OLSA-S 2 von $-7,43 \pm 0,55$ dB (Abb. 3). Es konnte kein signifikanter Unterschied des S/N50 abhängig von der Präsentationsreihenfolge des OLSA-S nachgewiesen werden ($t = 0,46$; $p = 0,65$).

Die mittleren S/N50 ergaben bei der Gruppe aFBE-S 1 $-1,74 \pm 0,63$ dB und bei der Gruppe aFBE-S 2 $-1,95 \pm 0,74$ dB (Abb. 4). Hinsichtlich der Präsentationsreihenfolge ließ sich auch beim aFBE-S kein signifikanter Unterschied des S/N50 nachweisen ($t = 0,93$; $p = 0,36$).

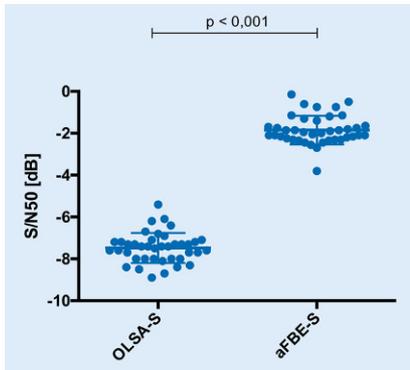


Abb. 5 ▲ Vergleich der Signal-Rausch-Abstände für das 50 %-Sprachverstehen in Dezibel (S/N50 [dB]) zwischen Oldenburger Satztest (OLSA-S) und dem adaptiven Freiburger Einsilbertest (aFBE-S) im Störschall

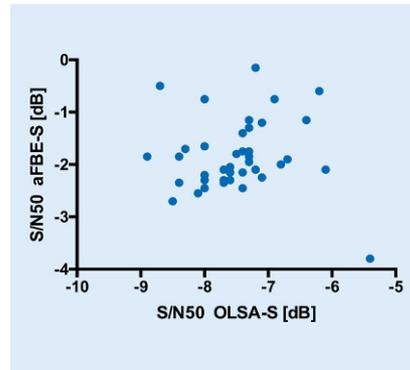


Abb. 6 ▲ Streudiagramm der Signal-Rausch-Abstände für das 50 %-Sprachverstehen des adaptiven Freiburger Einsilbertests im Störschall in Dezibel (S/N50 aFBE-S [dB]) und des Oldenburger Satztests im Störschall in Dezibel (S/N50 OLSA-S [dB])

Signal-Rausch-Abstände für das 50 %-Sprachverstehen und Korrelation

Der kleinste Signal-Rausch-Abstand für das 50 %-Sprachverstehen betrug beim OLSA $-8,9$ dB, während es beim aFBE $-3,8$ dB waren. Dem gegenüber waren die größten S/N50 beim OLSA $-5,4$ dB und beim aFBE $-0,15$ dB (■ **Abb. 5**). Der mittlere S/N50 beim OLSA war mit $-7,48 \pm 0,72$ dB kleiner als bei Anwendung des aFBE mit $-1,84 \pm 0,69$ dB. Es konnte ein signifikanter Unterschied der S/N50 der verwendeten Tests nachgewiesen werden ($t = 35,49$; $p < 0,001$).

Mit einem Korrelationskoeffizienten von $r = -0,03$ bei 40 Messungen und einem p -Wert von $p = 0,87$ konnte mittels Pearson-Korrelation zwischen den Ergebnissen des OLSA-S und des aFBE-S keine Korrelation gezeigt werden (■ **Abb. 6**).

Diskussion

Ziel dieser Arbeit war es, einen adaptiven Messalgorithmus auf den FBE im Störschall zu übertragen und diesen mit dem ebenfalls adaptiv messenden OLSA zu vergleichen. Dabei sollten die Ergebnisse hinsichtlich des Zeitbedarfs, möglicher geschlechtsspezifischer Unterschiede, der Präsentationsreihenfolge und auf eine Korrelation hin untersucht werden. Aufgrund der besonderen Bedeutung für die Hörgeräteversorgung

[12] wurden beide Sprachverstehenstests im Störschall angewendet.

Zeitbedarf und geschlechtsspezifische Unterschiede

Der für den aFBE-S gemessene Zeitbedarf unterscheidet sich von den Angaben in der Literatur zum klassischen FBE. So wird der Messzeitaufwand für eine Testliste des FBE mit 1,5 min beziffert [22]. Die Zeitdauer einer Schwellenbestimmung mittels 13 Listen und 3–4 unterschiedlichen Sprachpegeln wird mit 20 min (1200 s) angegeben (Genauigkeit 2 dB) [22]. Bei dem in dieser Studie durchgeführten adaptiven Messverfahren wurden 7,5 Testlisten dargeboten und mittels 30 Sprachpegeln präsentiert (der 31. Sprachpegel wurde aus dem vorherigen gemäß Algorithmus berechnet). Die durchschnittliche Dauer betrug dabei 518 ± 31 s. Demnach ist eine Schwellenbestimmung für den FBE-S mittels des verwendeten adaptiven Verfahrens schneller durchführbar als in der Literatur angegeben, obwohl die Einsilber bei bis zu 10-mal so häufigen verschiedenen Sprachpegeln präsentiert wurden. Im Vergleich mit dem OLSA-S zeigte der aFBE-S jedoch einen größeren Zeitbedarf. Die Anwendung des aFBE-S ließe sich zwar durch Gebrauch einer automatischen Anwendungssoftware wie der des OLSA beschleunigen. Ungeachtet dessen erfordert eine durchschnittliche adaptive Messung des FBE einen höheren

Zeitaufwand als die bisherige prozentuale Messweise bei 3 verschiedenen Sprachpegeln mit jeweils einer Testliste [22].

Der Messzeitaufwand für eine Testliste des OLSA wird mit 4 min (240 s) angegeben [22]. Mit einem in dieser Arbeit durchschnittlichen Zeitbedarf von 467 ± 37 s für die Messung von 2 Listen mit je 30 Sätzen entsprechen die erhobenen Ergebnisse der Literatur. Jedoch ist nicht erläutert, ob sich der angegebene Zeitaufwand von 4 min auf eine oder mehrere Trainingslisten mit 20 oder 30 Sätzen bezieht.

Hinsichtlich der Auswirkung des Geschlechts auf das Sprachverstehen konnte auch bei der adaptiven Messweise des FBE, wie zuvor in weiteren Studien beschrieben [24, 27], kein signifikanter geschlechtsspezifischer Unterschied der Ergebnisse nachgewiesen werden.

Einfluss der Präsentationsreihenfolge

Eine Erklärung für die fehlende Abhängigkeit des S/N50 von der Präsentationsreihenfolge könnte die in der Diskussion um einen möglichen Trainingseffekt des FBE-S angeführte Gewöhnung an die Testsituation sein [29]. Diese Habituation ist im Aufbau des adaptiven Algorithmus begründet. Die ersten 11 dargebotenen 5er-Gruppen des aFBE-S wirkten sich zwar auf den 12. präsentierten Sprachpegel aus, wurden jedoch nicht für die Berechnung der S/N50 berücksichtigt. Dadurch wurden die Probanden sowohl mit den Einsilbern und deren Präsentation als auch mit dem Testprozedere bereits vertraut gemacht. Für die Berechnung des Ergebnisses wurden lediglich die Sprachpegel 12–31, somit die letzten 20 Messungen, einbezogen. Demnach standen zum Ausgleich des bekannten prozeduralen Lerneffekts des OLSA-S, neben der Trainingsliste mit 30 Sätzen, weitere 11 Sätze aus der eigentlichen Testliste zur Verfügung.

Vergleich der Signal-Rausch-Abstände für das 50%-Sprachverstehen

Bei Normalhörenden ist der S/N50 i. Allg. kleiner als 0 dB und sowohl von dem individuellen Hörvermögen, dem präsentierten Sprachmaterial als auch dem dargebotenen Störschall abhängig [32]. Dadurch ergeben sich für die adaptive Messweise bei verschiedenen Sprachverstehenstests jeweils unterschiedliche Normwerte.

Im geschlechtsunabhängigen Vergleich der für den aFBE-S und den OLSA-S gefundenen Signal-Rausch-Abstände für das 50%-Sprachverstehen konnte ein signifikanter Unterschied nachgewiesen werden. So war der mittlere S/N50 des OLSA-S kleiner als der des aFBE-S. Dies könnte zum einen auf die unterschiedlichen Störgeräusche der beiden Sprachverstehenstests zurückzuführen sein. Während beim OLSA-S ein sprachsimulierendes Rauschen zur Anwendung kam, wurde beim aFBE-S das sog. CCITT-Rauschen mit einer unterschiedlichen Frequenzzusammensetzung verwendet. Zum anderen kann der Unterschied zwischen den S/N50 mit der Anwendung von Einsilbern erklärt werden, welche schwerer verständlich als ganze Sätze sind und dadurch eine Verschiebung der Diskriminationsfunktion mit höheren Sprachpegeln aufweisen [15]. Mehrsilbige und durch eine definierte Syntax miteinander in Verbindung stehende Wörter wie im OLSA-S sind dagegen durch eine höhere Redundanz besser zu verstehen [30].

Zudem können die verschiedenen Präsentationsweisen trotz Anwendung des gleichen Algorithmus zu unterschiedlichen S/N50 bei normalhörenden Probanden geführt haben. Während beim aFBE-S jeder Einsilber nachgesprochen wurde, können beim OLSA-S auch kognitive Funktionen durch das Nachsprechen ganzer Sätze getestet worden sein. Darüber hinaus unterschieden sich die Präsentationsweisen durch die Verwendung einer zusätzlichen Trainingsliste für den OLSA-S.

In Anbetracht der phonemischen Unausgewogenheit des FBE wurde eine adaptive Pegelsteuerung bisher als nicht

sinnvoll erachtet [14]. Zudem wurde schon 1953 von Hahlbrock angeführt, dass die Messung mit halben Testlisten ungenau sei, da nur ganze Listen phonetisch ausgewogen seien [13]. Auf Listenebene zeigte sich kürzlich beim FBE entgegen der Kritik eine relativ geringe Abweichung von der Phonemverteilung der deutschen Sprache [9]. Dies gilt jedoch nicht für die Anwendung von halben Testlisten oder einer weiteren Unterteilung wie die in dieser Studie durchgeführten 5er-Gruppen. Dadurch könnte die für ganze Testlisten geltende relative phonemische Ausgewogenheit aufgehoben worden sein und sich auf diese Weise auf das Ergebnis ausgewirkt haben. Zudem spricht gegen die angewandte adaptive Messung des FBE-S die beschränkte Anzahl an Testlisten, welche eine Wiederholung der Einsilber nach 2-maliger Anwendung des aFBE-S zur Folge hätten, und der erhöhte Zeitbedarf. Andererseits wiesen die S/N50 des OLSA-S als auch die des FBE-S bei Normalhörenden eine vergleichsweise kleine Streuung auf. Allerdings ist die langfristige Messgenauigkeit durch den prozeduralen Lerneffekt insbesondere beim OLSA-S zu bezweifeln [28]. Angesichts der, wie in **Abb. 1** exemplarisch dargestellten, kleinen Streuung der Sprachpegel im Verlauf der Messung des aFBE-S ist zu diskutieren, ob eine geringere Anzahl an Einsilbern ausgereicht hätte, um den S/N50 zu ermitteln. Dadurch würde sich auch der Zeitaufwand einer Messung des aFBE-S verringern.

Bisher sind keine Umrechnungsfaktoren entwickelt worden, um die unterschiedlichen Sprachverstehenstests direkt miteinander vergleichen zu können [15]. Dies liegt zum einen an dem verschiedenartigen Aufbau von Wort- und Satztests sowie an den unterschiedlichen Messweisen z. B. zwischen dem adaptiven OLSA und dem nichtadaptiven FBE. Zum anderen war die Entwicklung von Umrechnungsfaktoren bis zu den vergleichenden Messungen zwischen dem nichtadaptiven FBE und anderen Sprachverstehenstests nicht notwendig [20, 31]. Um wie aus den vorherigen Vergleichen eine höhere Genauigkeit oder Testäquivalenz abzuleiten, sollten jedoch die unterschiedlichen Messme-

thoden vergleichbar sein. Aufgrund dessen wurde in dieser Arbeit der gleiche adaptive Messalgorithmus angewendet.

Die Umrechenbarkeit von Ergebnissen aus Wort- und Satztests ist zu bezweifeln: So zeigte sich zwischen dem adaptiven finnischen Matrixtest und dem bisherigen finnischen Worttest nur eine schwache Korrelation ($r = -0,38$; $p < 0,001$) [4]. Beim Vergleich des deutschen Matrixtests OLSA mit dem FBE war nur bei Berücksichtigung des besten Quartils eine Korrelation nachweisbar, ansonsten zeigte sich keine Korrelation ($r = -0,18$; $p = 0,13$) [20]. Dies könnte auf die verschiedenen Störgeräusche oder die unterschiedliche Messweise zurückzuführen sein. In dieser Studie zeigten die Ergebnisse des OLSA und des aFBE bei Anwendung des gleichen Algorithmus im Störschall ebenfalls keine Korrelation ($r = -0,03$; $p = 0,87$), sodass hierfür nicht die unterschiedlichen Messweisen, sondern die verschiedenen Störgeräusche und die spezifischen Diskriminationsfunktionen ursächlich erscheinen. Jedoch ist nicht auszuschließen, dass die unterschiedlichen Ergebnisse bei Normalhörenden allein durch die Messreliabilität zu erklären sind [35]. Aus diesem Grund ist die Untersuchung von Schwerhörenden notwendig, um eine mögliche fehlende Korrelation der S/N50 beider Sprachverstehenstests nachweisen zu können.

Auch bei Anwendung eines gleichen Algorithmus für die adaptive Sprachpegeländerung unterschieden sich der aFBE-S und der OLSA-S in dieser Studie im verwendeten Störschall, der Präsentationsweise und der Redundanz des Sprachmaterials, wodurch ein Vergleich beider Sprachverstehenstests nur bedingt möglich war. Durch die weiterhin eingeschränkte Vergleichbarkeit von Worttests, wie dem bisherigen FBE-S, und Satztests, wie dem OLSA-S, haben beide Sprachverstehenstests je nach Fragestellung ihre Berechtigung.

Fazit für die Praxis

- Die bei Satztests bereits etablierte adaptive Sprachpegeländerung konnte als Messverfahren auf den FBE im Störschall übertragen werden.

- Eine mögliche fehlende Korrelation zwischen den Signal-Rausch-Abständen für das 50%-Sprachverstehen von Einsilber- und Satztests sollte weiter untersucht werden.
- Einsilber- und Satztests sind nach wie vor nur bedingt miteinander vergleichbar.
- Zukünftig werden in der Sprachaudiometrie weiterhin sowohl adaptive Satztests, wie z. B. der GÖSA und der OLSA, als auch Wort- und Zahlenverstehenstests, wie z. B. der Freiburger Sprachtest ohne und mit Störschall, abhängig von der Fragestellung angewendet werden.

Korrespondenzadresse

PD Dr. med. habil. J. Löhler

Wissenschaftliches Institut für angewandte HNO-Heilkunde des Deutschen Berufsverbandes der HNO-Ärzte e. V. (WIAHNO) Maienbeek 1, 24576 Bad Bramstedt, Deutschland
loehler@hno-aerzte.de

Einhaltung ethischer Richtlinien

Interessenkonflikt. T. Memmeler, R. Schönweiler, B. Wollenberg und J. Löhler geben an, dass kein Interessenkonflikt besteht.

Alle im vorliegenden Manuskript beschriebenen Untersuchungen am Menschen wurden mit Zustimmung der zuständigen Ethik-Kommission, im Einklang mit nationalem Recht sowie gemäß der Deklaration von Helsinki von 1975 (in der aktuellen, überarbeiteten Fassung) durchgeführt. Von allen beteiligten Patienten liegt eine Einverständniserklärung vor.

Open Access. Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Literatur

1. Baljić I, Hoppe U (2016) Der Freiburger Einsilbertest auf dem Prüfstand. HNO 64:538–539
2. Baljić I, Winkler A, Schmidt T, Holube I (2016) Untersuchungen zur perceptiven Äquivalenz der Testlisten im Freiburger Einsilbertest. HNO 64:572–583

3. D'Agostino RB, Belanger A, D'Agostino RB Jr (1990) A suggestion for using powerful and informative tests of normality. Am Stat 44:316–321
4. Dietz A, Buschermöhle M, Sivonen V et al (2015) Characteristics and international comparability of the Finnish matrix sentence test in cochlear implant recipients. Int J Audiol 54:80–87
5. Deutsches Institut für Normung (1995) Tonträger mit Sprache für Gehörprüfung – Teil 1: Tonträger mit Wörtern nach DIN 45621-1 (Aufnahme 1969). DIN 45626-1. Beuth, Berlin
6. Deutsches Institut für Normung (2011) Akustik – Audiometrische Prüfverfahren – Teil 1: Grundlegende Verfahren der Luft- und Knochenleitungsschwelenaudiometrie mit reinen Tönen. DIN EN ISO 8253-1. Beuth, Berlin
7. Deutsches Institut für Normung (2010) Akustik – Audiometrische Prüfverfahren – Teil 2: Schallfeld-Audiometrie mit reinen Tönen und schmalbandigen Prüfsignalen. DIN EN ISO 8253-2. Beuth, Berlin
8. Deutsches Institut für Normung (2012) Akustik – Audiometrische Prüfverfahren – Teil 3: Sprachaudiometrie. DIN EN ISO 8253-3. Beuth, Berlin
9. Exter M, Winkler A, Holube I (2016) Phonemische Ausgewogenheit des Freiburger Einsilbertests. HNO 64:557–563
10. Franz S, Rosanowski F, Eysholdt U, Hoppe U (2011) Freiburger Sprachverständnistest bei jugendlichen Sprachheilschülern. Sprache Stimme Gehör 35:105–110
11. https://www.g-ba.de/downloads/62-492-1666/HilfsM-RL_2018-07-19_IK-2018-10-03.pdf. Zugegriffen: 30.11.2018
12. https://www.g-ba.de/downloads/40-268-4059/2016-11-24_HilfsM-RL_Freiburger-Einsilbertest_TrG.pdf. Zugegriffen: 23.06.2018
13. Hahlbrock KH (1953) Über Sprachaudiometrie und neue Wörterteste. Arch Ohren Nasen Kehlkopfheilkd 162:394–431
14. Hey M, Brademann G, Ambrosch P (2016) Der Freiburger Einsilbertest in der postoperativen CI-Diagnostik. HNO 64:601–607
15. Hoth S (2016) Der Freiburger Sprachtest. Eine Säule der Sprachaudiometrie im deutschsprachigen Raum. HNO 64:540–548
16. http://www.hoertech.de/images/hoertech/pdf/mp/produkte/olsa/HT.OLSA_Handbuch_Rev01_0_mitUmschlag.pdf. Zugegriffen: 23.06.2018
17. <http://www.itu.int/rec/T-REC-G.227-198811-1/en>. Zugegriffen: 23.06.2018
18. Hyde JS (2016) Sex and cognition: gender and cognitive functions. Curr Opin Neurobiol 38:53–56
19. Kießling J (2010) Moderne Sprachtests auf dem Weg von der Forschung in die Praxis. HNO 58:595–596
20. Knief A, Schmidt CM, Deuster D et al (2010) Vergleich des Freiburger Einsilbertests und des Oldenburger Satztests bei Cochlea-Implantat-Trägern. 27. Wissenschaftliche Jahrestagung der DGPP, Aachen
21. Kollmeier B, Lenarz T, Kießling J et al (2014) Zur Diskussion um den Freiburger Einsilbertest. HNO 62:49–53
22. Kollmeier B, Lenarz T, Winkler A et al (2011) Hörgeräteindikation und -überprüfung nach modernen Verfahren der Sprachaudiometrie im Deutschen. HNO 59:1012–1021
23. Kompis M, Krebs M, Häusler R (2005) Überprüfung der Bezugskurven der Schweizer Version des Freiburger Zahlen- und Einsilbertests. HNO 54:445–450
24. Löhler J, Akcicek B, Pilnik M et al (2013) Evaluation des Freiburger Einsilbertests im Störschall. HNO 61:586–591
25. Löhler J, Akcicek B, Wollenberg B et al (2014) Results in using the Freiburger monosyllabic speech test in noise without and with hearing aids. Eur Arch Otorhinolaryngol 272:2135–2142
26. Mallinger E (2011) Trainingseffekte und Listenäquivalenz des Freiburger Einsilbertests im Störschall. Dissertation. Friedrich-Alexander-Universität, Erlangen, Nürnberg
27. Ptok M, Lichte C, Buller N et al (2005) Ist die Lautdiskriminationsfähigkeit geschlechtsabhängig? Laryngorhinootologie 84:20–23
28. Schlüter A, Holube I, Lemke U (2012) Trainingseffekte bei normaler und schneller Sprache. 15. Jahrestagung der Deutschen Gesellschaft für Audiologie, Erlangen
29. Schmidt T, Baljić I (2016) Untersuchung zum Trainingseffekt des Freiburger Einsilbertests. HNO 64:584–588
30. Steffens T (2017) Die systematische Auswahl von sprachaudiometrischen Verfahren. HNO 65:219–227
31. Sukowski H, Brand T, Wagener KC, 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. HNO 58:597–604
32. Thiele C, Sukowski H, Lenarz T, Lesinski-Schiedat A (2012) Göttinger Satztest im Störgeräusch für verschiedene Gruppen von Schwerhörigkeit. Laryngorhinootologie 91:782–788
33. Wagener K, Brand T, Kollmeier B (1999) Entwicklung und Evaluation eines Satztests für die deutsche Sprache Teil III: Evaluation des Oldenburger Satztests. Z Audiol 38:86–95
34. Wagener K, Kühnel V, Kollmeier B (1999) Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests. Z Audiol 38:4–15
35. Winkler A, Holube I (2016) Test-Retest-Reliabilität des Freiburger Einsilbertests. HNO 64:564–571