

Audiovisuelle Realisierung des subjektiven Höranstrengungsmessverfahrens ACALES

Zusammenfassung

Das Sprachverstehen kann durch die Kombination von auditiven und visuellen Sprachmerkmalen positiv beeinflusst werden. Dabei profitieren nicht nur Personen mit einer Hörbeeinträchtigung von den zusätzlichen Informationen durch das Mundbild, sondern auch Normalhörende. Hinsichtlich der Höranstrengung ergaben sich bisher jedoch unterschiedliche Ergebnisse. So zeigt eine Studie eine Abnahme, eine andere hingegen eine Zunahme der Höranstrengung bei zusätzlicher Präsentation des Mundbildes. Ziel der vorliegenden Studie war die Messung der subjektiven Höranstrengung bei audiovisueller Darbietung im Vergleich zur rein akustischen Präsentation. Als Messverfahren wurde die adaptive Skalierungsmethode ACALES (Adaptive Categorical Listening Effort Scaling, siehe Krueger et al., J Acoust Soc Am. 2017 06;141(6):4680–93) verwendet. Da ACALES bisher nur für rein akustische Stimuli genutzt wurde, war vorab eine Erweiterung des Verfahrens um die Wiedergabe audiovisueller Stimuli notwendig. Sowohl in der akustischen als auch in der audiovisuellen Kondition wurden Sätze des Oldenburger Satztestes (OLSA) bei Anwesenheit verschiedener Hintergrundgeräusche dargeboten. Zusätzlich wurde in der audiovisuellen Kondition das entsprechende Mundbild der Sprecherin auf einem Bildschirm gezeigt. Die Messungen wurden mit 15 jungen Personen mit normalem Hörvermögen und zehn älteren Personen mit Hörbeeinträchtigung durchgeführt. Basierend auf den Messungen der Höranstrengung in der akustischen und audiovisuellen Kondition für verschiedene Maskierer wurden die intra- und interindividuelle Standardabweichung sowie die Test-Retest-Reliabilität bestimmt. Für beide Konditionen ergab sich eine Abhängigkeit vom Maskierer. Neben einem signifikanten Unterschied zwischen audiovisueller und akustischer Kondition wurde ein signifikanter Unterschied zwischen den Probandengruppen ermittelt. Das Messverfahren eignet sich zur Erfassung interindividueller Unterschiede in der Bewertung der Höranstrengung und weist eine gute Reliabilität auf.

Schlüsselwörter: Höranstrengung, ACALES, audiovisuell, Hörbeeinträchtigung, Sprache im Störgeräusch

Einleitung

Eine Beteiligung an Gesprächen ermöglicht die Teilhabe am sozialen Leben. Auch wenn das Gesprochene für den Zuhörer ausreichend laut und auch verständlich ist, kann die Situation dennoch als ermüdend und anstrengend empfunden werden. Dies zeigt sich sowohl bei Normalhörenden als auch bei Personen mit Hörbeeinträchtigung (im Folgenden als Schwerhörende bezeichnet) besonders in Umgebungen, welche von viel Hintergrundgeräusch oder Nachhall geprägt sind. Mit abnehmendem Signal-Rausch-Abstand (engl. signal-to-noise-ratio, SNR) sinkt das Sprachverstehen. Zeitgleich nehmen die Konzentration und der Aufwand, der zum Verstehen nötig ist, stetig zu [1]. Dieser zusätzliche Aufwand kann als Höranstrengung bezeichnet werden [2]. Neben dem Hintergrundge-

räusch (Maskierer) und damit einhergehend mit dem SNR kann Höranstrengung auch durch Faktoren wie Schwerhörigkeit [1], [3], [4] und Alter [5], [6], [7] negativ beeinflusst werden. Eine Versorgung mit Hörgeräten kann sich hingegen positiv auf die Höranstrengung auswirken [1], [3], [8], [9]. Ein weiterer bedeutender Einflussfaktor ist die Nutzung des Mundbildes. Die durch das Mundbild zusätzlich erhaltenen Informationen können genutzt werden, um nicht Verstandenes zu ergänzen und falsch Verstandenes zu korrigieren [10], [11]. Studien zeigten, dass das Sprachverstehen dadurch nachweislich verbessert werden kann [9], [12], [13]. Hingegen ist der Einfluss auf die Höranstrengung nicht eindeutig. Neben einer Abnahme der Höranstrengung [14] kann die Nutzung des Mundbildes auch zu einer Zunahme der Höranstrengung führen [5] oder unverändert sein [9].

Saskia Ibelings^{1,2}

Michael Schulte^{1,3}

Melanie Krüger^{1,3}

Inga Holube^{2,3}

1 Hörzentrum Oldenburg GmbH, Oldenburg, Deutschland

2 Institut für Hörtechnik und Audiologie, Jade Hochschule, Oldenburg, Deutschland

3 Exzellenzcluster „Hearing4all“, Oldenburg, Deutschland

Die Ergebnisse der genannten Studien wurden mit objektiven Messverfahren (wie z.B. Dual-Task Paradigmen) erzielt. In dieser Studie wurde im Gegensatz dazu ein adaptives Messverfahren zur Ermittlung der subjektiven Höranstrengung (engl. Adaptive CAtegorical Listening Effort Scaling, ACALES, [15]) verwendet. Bei ACALES werden Sätze des Oldenburger Satztests [16] im Störgeräusch dargeboten. Die Aufgabe der Probanden besteht darin, die wahrgenommene Höranstrengung auf einer 13-stufigen Kategorien-Skala von „müheless“ bis „extrem anstrengend“ und einer 14. zusätzlichen Kategorie „nur Störgeräusch“ zu bewerten. Jeder Kategorie wird dabei ein Zahlenwert von 1 ESCU („müheless“, Einheit: effort scaling categorical unit) bis 13 ESCU („extrem anstrengend“) bzw. 14 ESCU („nur Störgeräusch“) zugeordnet. Während der Messung wird je nach Antwort der Probanden der SNR adaptiv verändert. Als Ergebnis ergibt sich eine Funktion, bei der jeder Höranstrengungskategorie einem SNR-Wert zugeordnet wird [15].

Um den Einfluss des Mundbildes auf die subjektive Höranstrengung zu untersuchen, war eine Erweiterung des Messverfahrens ACALES, welches bisher nur für rein akustische Stimuli genutzt wurde, auf die Wiedergabe audiovisueller Stimuli notwendig. Basierend auf den Erkenntnissen aus vorherigen Studien wurde ein Einfluss des Mundbildes auf die Höranstrengung [5], [14] und ein Unterschied der Höranstrengungsbewertungen zwischen normal- und schwerhörenden Probanden [1], [3], [4] vermutet. Des Weiteren wurde erwartet, dass zeitlich stationäre Maskierer die Höranstrengung mehr erhöhen als zeitlich fluktuierende [1], [15]. Bei fluktuierenden Maskierern ist es möglich, in die zeitlichen Lücken zu hören, wodurch die Sprachverständlichkeit verbessert [17] und daher vermutlich die Höranstrengung reduziert wird. Außerdem weisen Maskierer, welche das gleiche Geschlecht wie der OLSA-Sprecher haben, ein ähnliches Frequenzspektrum auf. Daher wurde vermutet, dass diese Maskierer zu einer höheren Höranstrengung führen als Maskierer mit dem Spektrum eines andersgeschlechtlichen Sprechers [18]. Um das erweiterte Messverfahren zu evaluieren, wurden die intra- und interindividuellen Standardabweichungen sowie die Test-Retest-Reliabilität bestimmt. Für die Studie ergaben sich somit folgende Ziele:

- Ermittlung der subjektiven Höranstrengung
 - bei rein akustischen und audiovisuellen Stimuli
 - für verschiedene Maskierer
 - bei Normal- und Schwerhörenden
- Bestimmung der intra- und interindividuellen Standardabweichung
- Ermittlung der Test-Retest Reliabilität

Die Untersuchung beschränkt sich auf die Analyse der Höranstrengung. Obwohl auch der Zusammenhang zwischen Höranstrengung und Sprachverstehen in diesen Konditionen von Interesse gewesen wäre, wurde auf die gleichzeitige Bestimmung des Sprachverstehens aus Aufwandsgründen verzichtet. Da bekannt ist, dass visuelle Sprachmerkmale die Sprachverständlichkeit deutlich

verbessern können, wodurch wiederum die Höranstrengung beeinflusst wird, wurde diesem Umstand durch entsprechende Verweise an mehreren Stellen dieser Arbeit Rechnung getragen.

Material und Methodik

Probanden

Insgesamt nahmen 25 Probanden an den Messungen teil. Die erste Probandengruppe umfasste 15 normalhörende (NH) Probanden im Alter von 19 bis 27 Jahren ($22,7 \pm 2,3$ Jahre). Davon waren sieben männlich und acht weiblich. Normalhörigkeit wurde in Anlehnung an Krueger et al. [15] durch einen PTA_4 (gemittelte Luftleitungsschwellen bei 500 Hz, 1000 Hz, 2000 Hz und 4000 Hz) kleiner als 20 dB HL definiert. Der gemittelte PTA_4 betrug auf dem rechten Ohr $2,0 \pm 2,8$ dB HL und auf dem linken Ohr $2,3 \pm 3,2$ dB HL. Zwei Probanden hatten bereits Erfahrung mit dem Messverfahren ACALES. Ein Proband kannte die OLSA-Stimuli nicht.

Zehn Schwerhörende im Alter von 63 bis 76 Jahren ($70,1 \pm 4,2$ Jahre) bildeten die zweite Probandengruppe. Sechs der Probanden waren männlich, vier weiblich. Ihr gemittelter PTA_4 betrug $43,1 \pm 5,3$ dB HL auf dem rechten und $40,3 \pm 5,3$ dB HL auf dem linken Ohr. Drei der Probanden waren mit Hörgeräten versorgt, alle Messungen wurden jedoch unversorgt durchgeführt. Den meisten Probanden war der OLSA bekannt, keiner der Probanden hatte Erfahrung mit der Ermittlung von Höranstrengung.

Stimuli

Als Stimuli wurden sowohl in der akustischen als auch in der audiovisuellen Kondition Sätze des OLSA mit einer weiblichen Sprecherin [19] genutzt. Bei den Sätzen des OLSA handelt es sich um Sätze mit der gleichbleibenden Reihenfolge von Wortkategorien Name-Verb-Zahl-Adjektiv-Objekt wie z.B. „Britta verleiht elf alte Bilder“. Jede Wortkategorie umfasst dabei zehn verschiedene Worte [16]. In der audiovisuellen Kondition wurden ergänzend die entsprechenden Mundbewegungen der Sprecherin auf einem Bildschirm gezeigt. Dazu wurde die Sprecherin, die auch den weiblichen OLSA eingesprochen hatte, nachträglich gefilmt [20]. Sie war ab dem Schulterbereich vor einem Greenscreen als Hintergrund sichtbar. In beiden Konditionen wurden zusätzlich verschiedene Maskierer wiedergegeben.

Maskierer

Die Messungen wurden insgesamt für drei verschiedene Maskierer durchgeführt. Neben dem weiblichen Olnoise (Olnoise mit dem Spektrum einer weiblichen Sprecherin [16], [19]) wurden der International Female Fluctuation Masker (IFFM; [21]) und OLSA-Sätze (männlich) [16] als Störsprecher genutzt. Beim Olnoise handelt es sich um ein stationäres Störgeräusch, welches durch mehrmalige

zufällige Überlagerungen der OLSA-Sätze generiert wurde [18]. Dadurch stimmt das gemittelte Langzeitspektrum des Olnoise mit dem der Stimuli überein. Beim IFFM handelt es sich um eine Variante des International Speech Test Signal (ISTS, [21]). Die Pausen wurden auf 250 ms gekürzt [22]. Für das ISTS bzw. den IFFM wurden Aufnahmen von sechs weiblichen Sprecherinnen unterschiedlicher Muttersprache (amerikanisches Englisch, Arabisch, Mandarin, Deutsch, Französisch und Spanisch) genutzt. Durch die Segmentierung und Mischung der Signale ist der Maskierer größtenteils unverständlich. Das gemittelte Langzeitspektrum entspricht weiblicher Sprache [21]. Für den Störsprecher wurden 15 Sätze des OLSA mit einem männlichen Sprecher [16] aneinandergereiht. Der Abstand zwischen den Sätzen betrug im Mittel ca. 500 ms. Es gab folglich eine zeitliche Lücke zwischen den Sätzen ohne Wortwiedergabe.

Apparatur

Die Messungen fanden in einem schallisolierten Raum statt. Die D/A-konvertierten Signale (Soundkarte ADI-8 PRO von RME; Haimhausen, München) wurden über einen Lautsprecher (Mackie HR 824; Bothell, Washington, USA), welcher in einem Abstand von ca. 1,3 m frontal zu den Probanden aufgestellt war, wiedergegeben. Des Weiteren befand sich vor den Probanden ein Touchscreen, auf dem die Skala zur Bewertung der Höranstrengung gezeigt wurde. Dieser Bildschirm wurde außerdem in der audiovisuellen Kondition zur Präsentation der Videos verwendet. Sowohl die akustischen als auch die audiovisuellen Stimuli wurden über den VLC-Mediaplayer (VideoLan, Version 3.0.3; Verden, Deutschland) dargeboten. Die Synchronisierung von Audio- und Videosignal wurde per Augenmaß durchgeführt. Die Höranstrengungsmessungen erfolgten mithilfe von ACALES [15], welches in MATLAB (Version 2007b; Natick, Massachusetts, USA) implementiert wurde.

Messverfahren

Zur Ermittlung der subjektiven Höranstrengung wurde ACALES [15] verwendet. Pro SNR wurden jeweils drei zufällige Sätze des OLSA wiedergegeben. Dabei wurde der Pegel der Zielsätze je nach Antwort der Probanden adaptiv verändert, der Pegel des Maskierers hingegen betrug konstant 65 dB SPL. Um zu laute bzw. zu leise Pegel zu verhindern, wurde der SNR-Bereich auf -35 bis 25 dB SPL beschränkt. Als Start-SNR wurden für die Normalhörenden 0 dB und für die Schwerhörenden 10 dB gewählt. Die Bewertung der subjektiven Höranstrengung erfolgte auf einer 14-stufigen Skala von „müheless“ (1 ESCU) bis „nur Störgeräusch“ (14 ESCU). In der ersten Phase von ACALES wurden die SNR-Werte für die Grenzen „müheless“ und „extrem anstrengend“ bestimmt. In der zweiten Phase wurden sieben verschiedene SNR für die benannten Kategorien in zufälliger Reihenfolge innerhalb des zuvor bestimmten SNR-Bereichs dargeboten. Nach einer Neuberechnung der Grenzen wurden in der dritten

Phase SNR-Werte für die sechs unbenannten Zwischenkategorien zweimal zufällig präsentiert. Insgesamt ergaben sich min. 21 SNR-Darbietungen pro Messung. Abbildung 1 stellt die verwendeten Skalen dar.

Da in der audiovisuellen Messkondition nicht auszuschließen war, dass die Probanden der Sprecherin aufgrund des Mundbildes folgen konnten, obwohl das Signal akustisch nicht wahrnehmbar war, wurde die Kategorie „nur Störgeräusch“ in „nichts wahrnehmbar“ umbenannt. Diese Kategorie bedeutet, dass die Sätze oder einzelne Bestandteile weder auditiv noch visuell über das Mundbild erschlossen werden können. Die Bedeutungen der Kategorien „nur Störgeräusch“ und „nichts wahrnehmbar“ wurde den Probanden vor den Messungen genau erklärt.

Messablauf

Die Studie wurde durch die Ethik-Kommission entsprechend des Ethik-Antrags Drs. 47/2017 freigegeben. Der Messablauf ist schematisch in Abbildung 2 dargestellt. Nach Aufklärung und schriftlicher Einwilligung der Probanden wurde die Luftleitungshörschwelle bestimmt. Anschließend wurde mit der Ermittlung der subjektiven Höranstrengung fortgefahren. Die Messungen der Schwerhörenden erfolgten ohne Hörgeräte. An einem Termin wurde die Höranstrengung in beiden Konditionen (akustisch und audiovisuell) jeweils für alle Maskierer ermittelt. Neben der Reihenfolge der Konditionen wurde auch die Reihenfolge der Maskierer innerhalb einer Kondition randomisiert. Bei der ersten Messkondition wurde für jeden der Maskierer ein Training vor der eigentlichen Messung durchgeführt, um die Stimuli, Maskierer sowie das Messverfahren kennenzulernen. In der zweiten Kondition waren Maskierer, Messverfahren sowie OLSA-Sätze bereits bekannt und es wurde lediglich ein Training zum Start durchgeführt, um sich mit der neuen Kondition (audiovisuell/akustisch) vertraut zu machen. Bei dem Training handelte es sich um die erste Messphase von ACALES. Die Stimuli wurden sowohl während des Trainings als auch während der Messung in zufälliger Reihenfolge präsentiert. Für zehn der 15 normalhörenden Probanden wurde an einem zweiten Termin erneut die Höranstrengung in beiden Konditionen mit allen Maskierern ermittelt. An diesem Termin wurde pro Kondition nur ein Training durchgeführt, da Messverfahren, Stimuli und Maskierer als bekannt vorausgesetzt wurden. Am ersten Termin wurden folglich je Proband vier Trainingseinheiten und sechs Messungen durchgeführt. Beim zweiten Termin wurde die Anzahl der Trainingseinheiten auf zwei reduziert.

Analyse und Statistik

Für alle Probanden wurden Höranstrengungskurven nach der in ACALES integrierten BX-Anpassungsmethode [23] erstellt. Krueger et al. [1], [15] konnten bereits zeigen, dass diese Anpassungsmethode i.d.R. zu einer validen Kurvenanpassung führt. Diese Kurven verlaufen jeweils von 1 ESCU bis 7 ESCU und von 7 ESCU bis 13 ESCU li-



Abbildung 1: Skalen zur Erfassung der subjektiven Höranstrengung

Links: Skala der akustischen Kondition, rechts: Skala für die audiovisuelle Kondition; die Kategorie „nur Störgeräusch“ wurde in „nichts wahrnehmbar“ umbenannt.

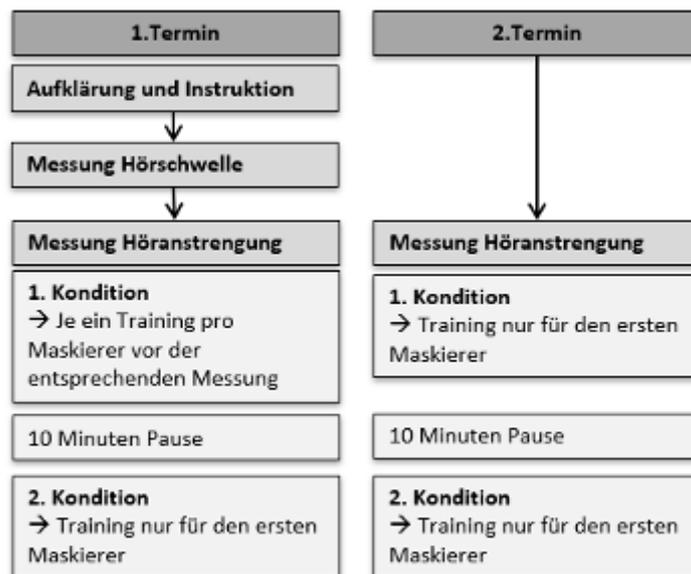


Abbildung 2: Schematische Darstellung des Messablaufs

Die Reihenfolge der Konditionen (akustisch und audiovisuell) sowie der Maskierer (Olnoise, IFFM, Störsprecher) wurden randomisiert. Zehn der 15 Normalhörenden nahmen an dem zweiten Termin teil.

near. Der Schnittpunkt der Geraden wurde zwischen 5 ESCU und 9 ESCU geglättet [15]. Die Höranstrengungskurven wurden über alle Probanden pro Kondition (d.h. Präsentationsmodus), Maskierer und Höranstrengungskategorie gemittelt.

Die Ergebnisse wurden mit SPSS 25.0.0 ausgewertet. Nach dem Kolmogorov-Smirnov-Test sind alle Daten normal verteilt. Ein normalhörender Proband wurde als Ausreißer von der Auswertung ausgeschlossen, da die Daten um mehr als die dreifache Standardabweichung

abwichen. Bei den Schwerhörenden musste ein Proband ausgeschlossen werden, da die Höranstrengung unabhängig von der Kondition und den Maskierern nie geringer als mit 4 ESCU bewertet wurde. Dadurch war keine valide Anpassung der Höranstrengungskurve möglich, denn die Höranstrengung dieses Probanden wurde folglich durch die mithilfe der Anpassungsmethode berechneten Werte nicht repräsentiert.

Als statistische Tests wurden Varianzanalysen (ANOVA) mit Messwiederholung und einem Signifikanzniveau von

$\alpha=0,05$ genutzt. Als Innersubjektfaktoren wurden stets die benannten Kategorien (1 ESCU, 3 ESCU, 5 ESCU, 7 ESCU, 9 ESCU, 11 ESCU und 13 ESCU) sowie die Maskierer (Olnoise, IFFM, Störsprecher) gewählt. Je nach Fragestellung gab es einen zusätzlichen Innersubjektfaktor Kondition (akustisch, audiovisuell) oder Zeit (1. Termin, 2. Termin) oder einen Zwischensubjektfaktor Probandengruppe (normalhörend, schwerhörend). Als Post-hoc Tests wurden t-Tests durchgeführt. Diese wurden aufgrund der linearen Verläufe der Höranstrengungsfunktionen für die Kategorien 1 ESCU, 7 ESCU und 13 ESCU durchgeführt. Pro Kondition wurden aufgrund der drei gewählten Kategorien und der drei verwendeten Maskierer jeweils neun t-Tests durchgeführt. Zum Vergleich der Konditionen und der Ergebnisse des ersten und zweiten Messtermins wurden aus diesem Grund ebenfalls je neun t-Tests durchgeführt. Daher betrug das nach Bonferroni korrigierte Signifikanzniveau in allen Fällen $\alpha=0,05/9=0,005$.

Ergebnisse

Höranstrengung in der akustischen Kondition

Bei den normalhörenden Probanden ist in der akustischen (a) Kondition eine Abhängigkeit der Höranstrengung vom Maskierer erkennbar (siehe Abbildung 3, oben). Eine Verschiebung der Kurven zu niedrigeren SNR-Werten bedeutet eine geringere Höranstrengung. Bei niedrigen SNR-Werten wurde das Olnoise als am anstrengendsten empfunden, während der Störsprecher als am wenigsten anstrengend wahrgenommen wurde. Eine ANOVA mit Messwiederholung bestätigte den Unterschied in der Bewertung der Maskierer [Greenhouse-Geisser $\epsilon=0,682$, $F(1,364, 17,735)=21,481$, $p<0,001$]. Außerdem wurde ein signifikanter Unterschied zwischen den Kategorien [$F(6, 78)=380,456$, $p<0,001$] gefunden. Auch eine Interaktion zwischen Maskierer und Kategorien wurde ermittelt [$F(12, 156)=77,175$, $p<0,001$]. Unter Berücksichtigung der Bonferroni-Korrektur ergaben t-Tests für verbundene Stichproben zwischen Olnoise und IFFM einen signifikanten Unterschied für die Kategorie „extrem anstrengend“ (13 ESCU). Olnoise und Störsprecher weisen zusätzlich in der Kategorie „mittelgradig anstrengend“ (7 ESCU) einen signifikanten Unterschied auf. Bei 7 ESCU unterscheiden sich auch IFFM und Störsprecher signifikant (alle $p<0,001$). In der Kategorie „müheless“ (1 ESCU) gibt es einen signifikanten Unterschied zwischen der Olnoise und IFFM ($p=0,018$).

Im Gegensatz dazu weisen die Höranstrengungskurven der Maskierer bei den Schwerhörenden in der akustischen Kondition kaum Unterschiede auf (siehe Abbildung 3, unten). Die ANOVA mit Messwiederholung ergab, dass sich zwar die Kategorien signifikant unterscheiden [$F(2,6)=3,292$, $p<0,001$], die Bewertungen der Maskierer weisen hingegen keine signifikanten Unterschiede auf

[$F(6,48)=143,908$, $p=0,063$]. Eine signifikante Interaktion zwischen Maskierer und Kategorie ist nachweisbar [$F(12,96)=2,710$, $p=0,003$]. Dies verdeutlicht, dass die Unterschiede in der Bewertung der Maskierer abhängig von der Höranstrengungskategorie sind. Unterschiede zeigen sich bei niedrigen SNR, also hohen Höranstrengungskategorien. Je niedriger der SNR-Bereich, desto höhere SNR-Unterschiede ergaben sich zwischen den Maskierern, um die gleiche Höranstrengungsbewertung zu erzielen.

Höranstrengung in der audiovisuellen Kondition

Auch in der audiovisuellen (av) Kondition ist bei den Normalhörenden eine unterschiedliche Bewertung der Maskierer ersichtlich (siehe Abbildung 3, oben). Beim Störsprecher wurden insgesamt die geringsten SNR-Werte erzielt, sodass dieser folglich als am wenigsten anstrengend wahrgenommen wurde, während das Olnoise aufgrund der höchsten SNR-Werte als am anstrengendsten empfunden wurde. Die Unterschiede zwischen den Maskierern nehmen mit steigender Höranstrengung zu. Zwischen Olnoise und Störsprecher beträgt die Differenz im SNR bei 13 ESCU ca. 14 dB. Beim IFFM sind im Mittel 2,5 dB höhere SNR-Werte als beim Störsprecher nötig. Eine ANOVA mit Messwiederholung zeigt eine signifikante unterschiedliche Bewertung der Maskierer [$F(2, 26)=15,774$, $p<0,001$] sowie einen signifikanten Unterschied der Kategorien [$F(6, 78)=334,995$, $p<0,001$]. Die Interaktion zwischen Maskierer und Kategorien ist ebenfalls signifikant [$F(12, 156)=25,674$, $p<0,001$]. Die t-Tests für abhängige Stichproben ergaben unter Berücksichtigung der Bonferroni-Korrektur signifikante Unterschiede zwischen Olnoise und IFFM bei 7 ESCU ($p=0,004$) und 13 ESCU ($p<0,001$) sowie zwischen Olnoise und Störsprecher ebenfalls in den Kategorien 7 ESCU und 13 ESCU (jeweils $p<0,001$). Für IFFM und Störsprecher konnten hingegen keine signifikanten Unterschiede ermittelt werden ($p>0,05$). Auch bei den Schwerhörenden wurde der Störsprecher tendenziell als am wenigsten anstrengend bewertet (siehe Abbildung 3, unten). Beim Vergleich zwischen IFFM und Störsprecher beträgt die Differenz bis 7 ESCU ca. 2 dB, bis 13 ESCU steigt die Differenz auf ungefähr 4 dB an. Zwischen Olnoise und IFFM ist kein deutlicher Unterschied ersichtlich. Die ANOVA ergab, dass die Bewertungen der Maskierer [$F(2,16)=6,393$, $p=0,009$] und die Kategorien [$F(6,48)=123,614$, $p<0,001$] sich signifikant unterscheiden. Zudem wurde eine signifikante Interaktion zwischen Kategorie und Maskierer festgestellt [$F(12,96)=8,640$, $p<0,001$]. Post-hoc t-Tests mit einem nach Bonferroni korrigierten Signifikanzniveau zeigten, dass sich die SNR-Werte von Störsprecher und Olnoise bei 13 ESCU ($p=0,002$) sowie Störsprecher und IFFM bei 7 ESCU ($p=0,004$) signifikant unterscheiden.

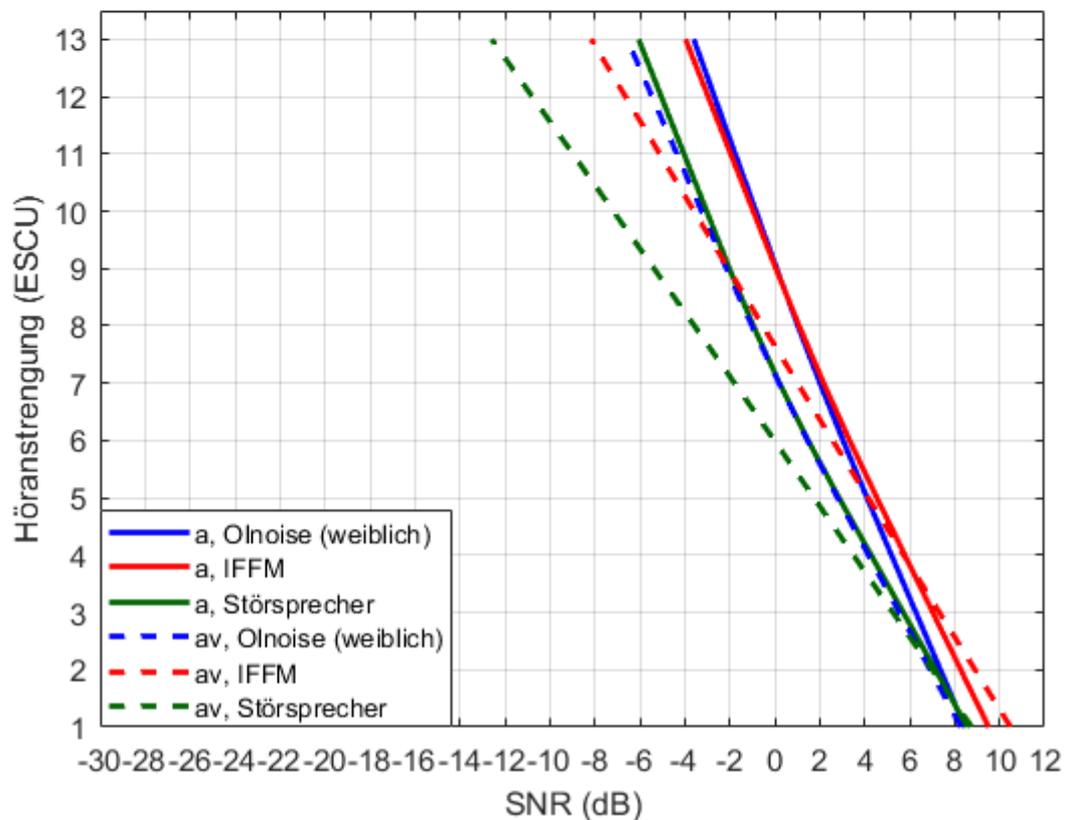
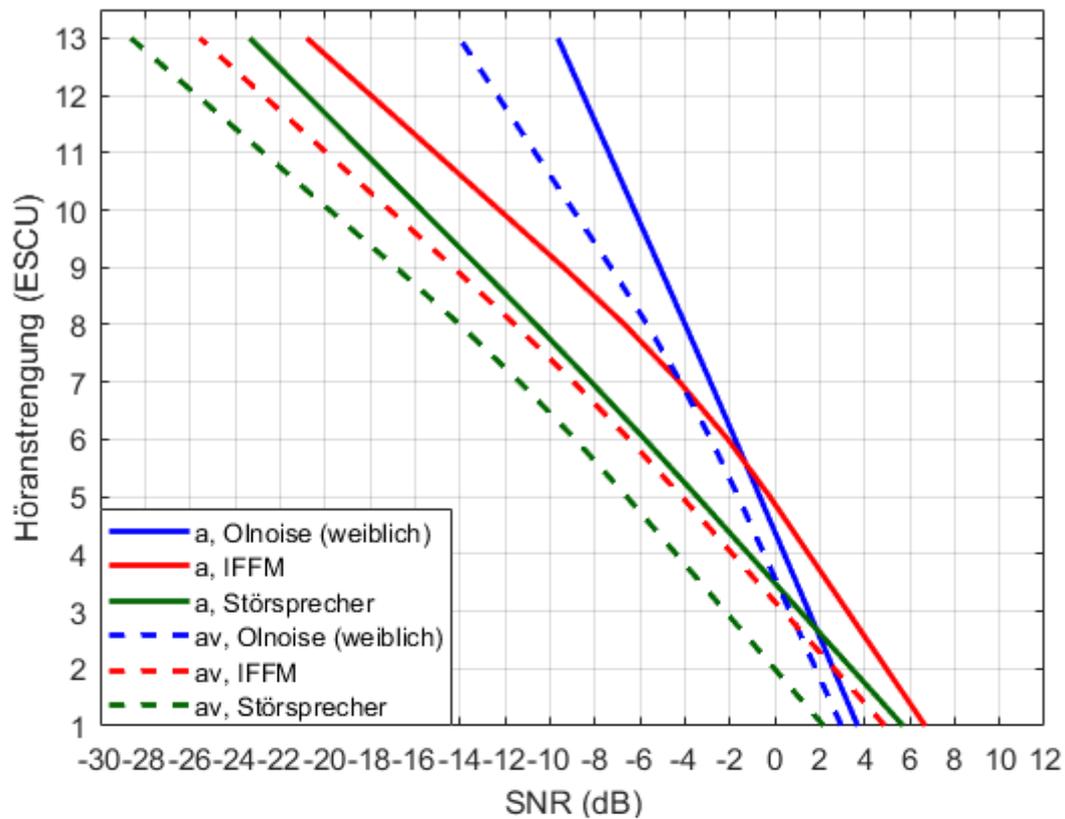


Abbildung 3: Oben: Gemittelte Höranstrengungskurven der Normalhörenden (n=14); unten: Gemittelte Höranstrengungskurven der Schwerhörenden (n=9)

Die durchgezogenen Linien repräsentieren jeweils die Ergebnisse der akustischen (a) Kondition und die gestrichelten Linien die Ergebnisse der audiovisuellen (av) Kondition.

Vergleich der Konditionen

Sowohl bei den Normalhörenden als auch bei den Schwerhörenden ist eine Verschiebung der Höranstrengungskurven der audiovisuellen Kondition relativ zur akustischen Kondition zu niedrigeren SNR-Werten ersichtlich (siehe Abbildung 3). Die audiovisuelle Kondition wurde somit als weniger anstrengend als die akustische Kondition empfunden. Bei den Normalhörenden sind die Kurven der Konditionen beim IFFM und Störsprecher nahezu parallel zueinander. Für die akustische Kondition sind bei diesen Maskierern im Mittel ca. 4 dB höhere SNR-Werte nötig, um die gleiche Höranstrengungsbewertung zu erzielen. Beim Olnoise beträgt die Differenz bis 5 ESCU ca. 1 dB und steigt auf 4 dB bei 13 ESCU an. Die ANOVA mit Messwiederholung bestätigte einen signifikanten Unterschied in der Bewertung der Konditionen [$F(1, 13)=14,656, p=0,002$]. Zusätzlich wurden signifikante Interaktionen zwischen Kategorie und Maskierer [$F(12, 156)=64,328, p<0,001$] und Kondition und Kategorie [$F(6,78)=3,630, p<0,003$] ermittelt. Die Post-hoc t-Tests für abhängige Stichproben (siehe Tabelle 1, Korrektur nach Bonferroni) ergaben signifikante Unterschiede der Konditionen beim IFFM bei 7 ESCU ($p=0,004$) und beim Störsprecher bei 13 ESCU ($p<0,001$). Beim Olnoise konnte hingegen kein signifikanter Einfluss des Mundbildes ermittelt werden.

Bei den Schwerhörenden nimmt der Unterschied der Konditionen unabhängig vom Maskierer mit steigender Höranstrengung zu. Während die Differenz der Konditionen bei 13 ESCU beim Olnoise lediglich 2 dB beträgt, liegt diese beim IFFM bereits bei 4 dB und beim Störsprecher bei 6 dB. Eine ANOVA mit Messwiederholung konnte jedoch keinen signifikanten Einfluss der Konditionen auf die Höranstrengungsbewertung ermitteln [$F(1,8)=4,760, p=0,061$]. Jedoch ergaben sich nicht nur signifikante Interaktionen zwischen Kondition und Kategorie sowie Maskierer und Kategorie (jeweils $p<0,001$) sondern auch zwischen Kondition, Maskierer und Kategorie ($p=0,012$). Aufgrund der signifikanten Interaktionen mit dem Faktor Kondition wurden t-Tests für verbundene Stichproben durchgeführt. Für alle Maskierer konnte bei 13 ESCU ein signifikanter Effekt bei der zusätzlichen visuellen Darbietung des Mundbildes gefunden werden (jeweils $p<0,01$ unter Berücksichtigung der Bonferroni-Korrektur).

Vergleich der Probandengruppen

Abbildung 4 stellt den Vergleich der Probandengruppen dar. Es ist erkennbar, dass die Schwerhörenden unabhängig von Kondition und Maskierer tendenziell bei höheren SNR-Werten ähnliche Höranstrengungsbewertungen wie die Normalhörenden zeigen. Während beim Olnoise die Kurven in beiden Konditionen nahezu konstant um ca. 5 dB verschoben sind, nimmt die Differenz bei den anderen Maskierern und Konditionen mit steigender Höranstrengung auf bis zu jeweils ca. 17 dB bei 13 ESCU zu. Die ANOVA mit Messwiederholung und dem zusätzlichen Zwischensubjektfaktor Probandengruppe (normalhörend,

schwerhörend) bestätigte einen signifikanten Unterschied zwischen den Probandengruppen. Des Weiteren ergaben sich signifikante Interaktionen zwischen Maskierer und Probandengruppe, Kategorie und Probandengruppe, Maskierer, Kategorie und Probandengruppe sowie Kondition, Maskierer, Kategorie und Probandengruppe (alle $p<0,05$). Die t-Tests für unabhängige Stichproben mit einem nach Bonferroni korrigiertem Signifikanzniveau belegten, dass sich die Bewertungen der Normal- und Schwerhörenden unabhängig von Kondition und Maskierer bei 7 ESCU und 13 ESCU signifikant unterscheiden. Für das Olnoise wurde in der akustischen Kondition bei 1 ESCU ebenfalls ein signifikanter Unterschied ermittelt (jeweils $p<0,05$ unter Berücksichtigung der Bonferroni-Korrektur).

Intra- und interindividuelle Standardabweichung

Die intra- und interindividuellen Standardabweichungen wurden für die Messergebnisse von zehn Normalhörenden bestimmt. Dazu wurden die SNR-Werte der benannten Kategorien für jede Messkondition und jeden Maskierer genutzt. Die intraindividuelle Standardabweichung wurde aus den SNR-Werten der beiden Messtermine für jeden Probanden berechnet. Die Ergebnisse für eine Höranstrengungskategorie, einen Maskierer und einer Darbietungskondition wurden durch Mittelung der intraindividuellen Standardabweichungen aller Probanden bestimmt. Bei der interindividuellen Standardabweichung wurde die Standardabweichung der SNR-Werte aller Probanden pro Messtermin für jede Kondition, jeden Maskierer und jede benannte Kategorie bestimmt und anschließend über die zwei Messtermine gemittelt. Tabelle 2 beinhaltet die berechneten Werte. Für die intraindividuelle Standardabweichung ergaben sich Werte zwischen 1,8 bis 2,6 dB und die interindividuelle Standardabweichung umfasst Werte von 4,3 bis 8,7 dB. Die Werte in der audiovisuellen Kondition sind bei beiden Standardabweichungen tendenziell höher als in der akustischen Kondition.

Test-Retest Reliabilität

In Abbildung 5 sind die Höranstrengungskurven der beiden Messtermine pro Kondition und Maskierer aufgetragen. Am zweiten Termin wurden im Mittel nahezu identische Ergebnisse erreicht. Zur Untersuchung der Test-Retest Reliabilität wurde eine ANOVA mit Messwiederholung und dem zusätzlichen Faktor Zeit (1. Termin, 2. Termin) durchgeführt. Dabei ergab sich kein signifikanter Unterschied zwischen den Ergebnissen der beiden Messtermine [$F(1,9)=0,205, p=0,662$]. Weiter konnten keine signifikanten Interaktionen mit dem Faktor Zeit nachgewiesen werden (alle $p>0,05$).

Zusätzlich wurde für jede Kondition und jeden Maskierer für die benannten Kategorien der Intraklassen-Korrelationskoeffizient (IKK) bestimmt (siehe Tabelle 3). Der gemittelte IKK der akustischen Kondition ist mit 0,856 etwas geringer als in der audiovisuellen Kondition (0,915).

Tabelle 1: Ergebnisse der Post-hoc Tests zum Vergleich der akustischen (a) und der audiovisuellen (av) Kondition für die Normalhörenden (NH) und Schwerhörenden (SH)

	Olnoise		IFFM		Störsprecher	
	NH	SH	NH	SH	NH	SH
1 ESCU	0,306	0,876	0,109	0,636	0,007	0,898
7 ESCU	0,021	0,040	0,004*	0,469	0,038	0,149
13 ESCU	0,009	< 0,001*	0,051	0,003*	< 0,001*	0,002*

* Signifikanter Unterschied bei Verwendung eines Signifikanzniveaus von $\alpha=0,005$

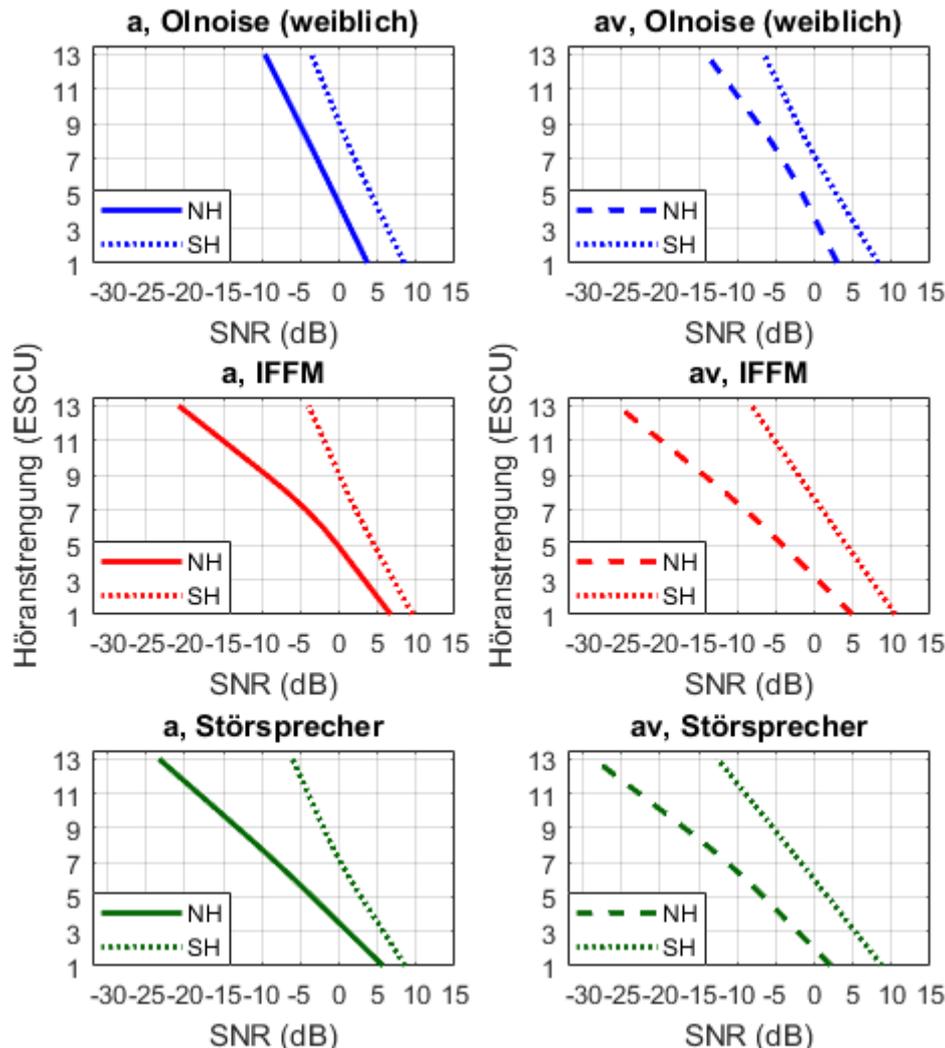


Abbildung 4: Vergleich der Höranstrengungskurven der Normalhörenden (n=14; gestrichelt) und Schwerhörenden (n=9; gepunktet) für alle Maskierer in der akustischen (a) und audiovisuellen (av) Kondition

Die geringsten Werte werden beim Olnoise sowohl in der akustischen (0,699) als auch in der audiovisuellen Kondition (0,717) jeweils bei 1 ESCU erreicht. Der höchste IKK ergibt sich für die audiovisuellen Kondition beim Olnoise bei 13 ESCU (0,983).

Diskussion

Vergleich der Maskierer

Die Ergebnisse zeigen, dass die Bewertung der Höranstrengung sowohl in der akustischen als auch in der au-

diavisuellen Kondition eine Abhängigkeit vom Maskierer aufweist. Signifikante Unterschiede waren jedoch erst bei zunehmender Höranstrengung nachweisbar. Dabei wurde das Olnoise als am anstrengendsten empfunden, während der Störsprecher als am geringsten anstrengend wahrgenommen wurde. Eine Erklärung der unterschiedlichen Wahrnehmung liegt in der zeitlichen und spektralen Struktur der Maskierer. Das Langzeitspektrum des Olnoise entspricht dem der Stimuli, sodass der Maskierungsgrad maximal ist [16]. Maskierer, welche das gleiche Geschlecht wie das Nutzsignal aufweisen, haben zudem eine höhere Maskierungswirkung [18]. Das IFFM entspricht im Langzeitspektrum ebenfalls weiblicher Sprache, ist

Tabelle 2: Intra- und interindividuelle Standardabweichung der akustischen (a) und der audiovisuellen (av) Kondition

		Intraindividuelle Standardabweichung in dB				Interindividuelle Standardabweichung in dB			
		Olnoise	IFFM	Störsprecher	Mittel	Olnoise	IFFM	Störsprecher	Mittel
1 ESCU	a	2,2	1,4	1,9	1,8	3,4	4,2	5,3	4,3
	av	3,1	2,8	1,9	2,6	5,4	6,7	7,8	6,6
3 ESCU	a	1,4	1,6	2,0	1,8	3,3	4,6	5,5	4,5
	av	2,9	2,7	1,7	2,4	6,1	7,3	8,1	7,2
5 ESCU	a	1,5	1,9	2,4	1,9	3,2	5,0	5,8	4,7
	av	2,6	2,6	2,1	2,4	6,9	7,9	8,6	7,8
7 ESCU	a	1,3	2,1	2,5	2,0	3,1	5,4	6,0	4,8
	av	2,4	2,5	2,5	2,5	7,6	8,4	8,8	8,3
9 ESCU	a	1,0	2,2	2,4	1,9	2,9	5,6	6,0	4,8
	av	2,0	2,2	2,7	2,3	8,1	8,7	8,9	8,6
11 ESCU	a	0,8	2,4	2,3	1,9	2,8	5,7	5,9	4,8
	av	1,7	1,9	2,9	2,2	8,4	8,7	8,7	8,6
13 ESCU	a	1,0	2,8	2,3	2,9	2,8	5,9	5,9	4,9
	av	1,7	2,0	3,4	2,4	8,8	8,8	8,5	8,7

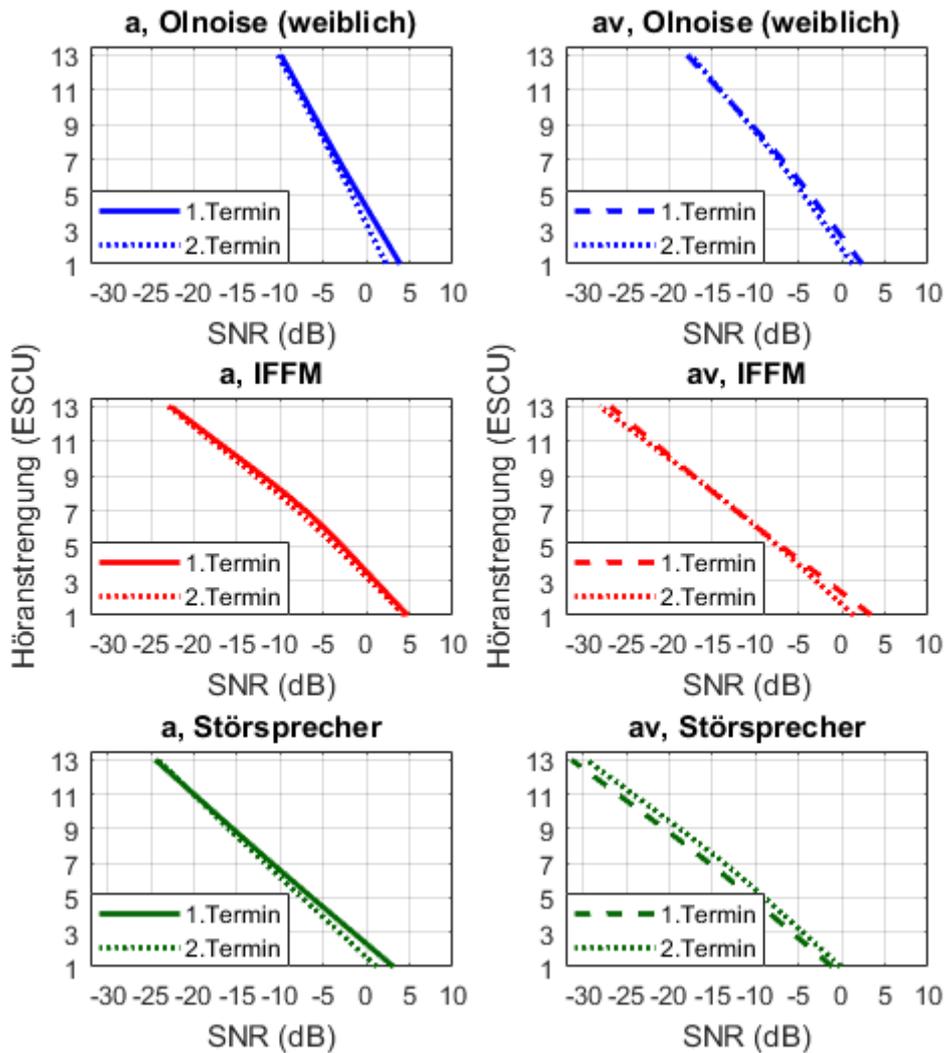


Abbildung 5: Gegenüberstellung der gemittelten Höranstrengungskurven (n=10 Normalhörende) für alle Maskierer in der akustischen (a) und audiovisuellen (av) Kondition für den 1. und 2. Messtermin

Tabelle 3: Intraklassen-Korrelationskoeffizienten für die akustische (a) und audiovisuelle (av) Kondition
Modell: zweifach, gemischt, Typ: absolut, durchschnittliche Maße

		Intraklassen-Korrelationskoeffizient							
		1 ESCU	3 ESCU	5 ESCU	7 ESCU	9 ESCU	11 ESCU	13 ESCU	Mittel
Olnoise	a	0,699*	0,747*	0,782*	0,806*	0,847*	0,892*	0,894*	0,810*
	av	0,717	0,851*	0,915*	0,948*	0,970*	0,982*	0,983*	0,909*
IFFM	a	0,893	0,921*	0,930*	0,924*	0,893*	0,848*	0,827*	0,891*
	av	0,821	0,849*	0,933*	0,951*	0,962*	0,966*	0,957*	0,920*
Störsprecher	a	0,885	0,906*	0,908*	0,901*	0,884*	0,836*	0,762*	0,869*
	av	0,941	0,959*	0,963*	0,956*	0,930*	0,874*	0,795*	0,917*

* Signifikant bei einem Signifikanzniveau von $\alpha=0,05$

jedoch fluktuierend [22], sodass es möglich ist, in die zeitlichen Lücken des Maskierers zu hören [17]. Der Störsprecher hingegen ist männlich, sodass die Maskierungswirkung dadurch bereits reduziert wird. Des Weiteren sind die Sprechpausen dieses Maskierers länger als die Lücken im IFFM. Infolgedessen sind Stimuli und Maskierer beim Störsprecher besser separierbar, sodass die subjektive Höranstrengung am geringsten ist. Insgesamt können alle genannten Faktoren einen Einfluss auf die Sprachverständlichkeit haben. Infolgedessen ist dadurch auch eine Beeinflussung der Höranstrengung möglich.

Eine Ausnahme davon bilden die Schwerhörenden in der akustischen Kondition, denn dort war kein Unterschied in der Bewertung der Höranstrengung zwischen den Maskierern feststellbar. Eine mögliche Erklärung liegt darin, dass die zeitliche und spektrale Auflösung aufgrund des Hörverlustes nur eingeschränkt oder gar nicht möglich ist [17].

Auch Krueger et al. [1], [15] stellten fest, dass stationäre Maskierer (Olnoise) als anstrengender empfunden werden als fluktuierende (IFFM und ISTS). Selbst bei vergleichbarer Sprachverständlichkeit waren die Unterschiede in der Höranstrengungsbewertung der Maskierer nachweisbar [1]. In einer anderen Studie ergab sich eine höhere Höranstrengung bei Nutzung eines Störsprechers als bei fluktuierendem oder stationärem Rauschen [24]. In dieser Studie von Koelewijn et al. [24] wurde kein subjektives Verfahren verwendet, sondern die Pupillometrie zur Ermittlung der objektiven Höranstrengung genutzt. Außerdem wurde der Störsprecher spektral verändert, sodass dieser Maskierer das gleiche Langzeitspektrum wie der Nutzsprecher aufwies. Da neben der energetischen Maskierung bei einem Störsprecher die informationelle Maskierung hinzukommt, wird die Maskierungswirkung erhöht und der Störsprecher als anstrengender bewertet. Fluktuierende und stationäre Störgeräusche erwiesen sich daher in dieser Studie als weniger anstrengend.

Vergleich der akustischen und der audiovisuellen Kondition

Wie erwartet, unterschieden sich die Höranstrengungsbewertungen der akustischen und der audiovisuellen Kondition bei beiden Probandengruppen signifikant. Die Nutzung des Mundbildes reduzierte die subjektive Höranstrengung. Der Vorteil durch das Mundbild zeigte sich jedoch erst bei negativen SNR-Werten, also bei höherer Höranstrengung. Bei höheren SNR-Werten ist die Sprachverständlichkeit hoch [1], sodass davon auszugehen ist, dass das Mundbild dort wenig Zusatzinformationen zur Verbesserung der Sprachverständlichkeit und Reduzierung der Höranstrengung liefert. Bei niedrigeren SNR-Werten bietet das Mundbild zusätzliche visuelle Informationen, sodass Nichtverstandenes ergänzt und Falschverstandenes korrigiert werden kann [11]. Des Weiteren können visuelle Sprachmerkmale zu einer besseren Separierung von Nutz- und Störsignal führen [12]. Durch beide Faktoren war vermutlich eine Zunahme der Sprachverständlichkeit möglich. Letztgenannter Faktor scheint gemäß den hier erzielten Ergebnissen für den Störsprecher einen größeren Effekt als beim IFFM zu haben.

Lorach et al. [20], welche dieselben Stimuli wie in dieser Studie verwenden, stellten bei Darbietung des Mundbildes eine Verbesserung der Schwelle für ein Sprachverstehen von 80% um ca. 4,5 dB SNR fest. Es ist insgesamt davon auszugehen, dass durch das verbesserte Sprachverstehen die Höranstrengung reduziert wurde [1]. Des Weiteren deutet der deutlich steilere Verlauf der Höranstrengungskurven für das Olnoise im Vergleich zum IFFM und zum Störsprecher auf einen Zusammenhang zwischen Sprachverstehen und Höranstrengung hin. Denn auch beim Testen der Sprachverständlichkeit ergibt sich für das Olnoise die steilste Funktion [1]. Auch Studien, welche den Einfluss von Störgeräusch und Nachhall bei Normal- und Schwerhörenden auf die Sprachverständlichkeit und die subjektive Höranstrengung untersucht haben, wiesen eine Abnahme der Höranstrengungen bei steigender Sprachverständlichkeit [25], [26] nach.

Neben der subjektiven Bewertung der Höranstrengung können auch objektive Verfahren verwendet werden.

Sommers und Phelps [14] untersuchten die Gedächtnisleistung bei jungen und älteren normalhörenden Probanden. Dabei wurden unterschiedlich lange Wortlisten in Ruhe verwendet, wobei die letzten drei Worte wiederholt werden sollten. Der Sprachpegel ist nicht bekannt. Es ist zu vermuten, dass dieser wie bei der von Sommers und Phelps [14] durchgeführten Sprachverständlichkeitsmessung 60 dB SPL betrug. Eine Zunahme der richtig wiederholten Worte in der audiovisuellen Kondition deutete dabei auf eine Abnahme der Höranstrengung. Im Gegensatz zu den jungen Probanden konnten bei den älteren Probanden, welche eine signifikant schlechtere Hörschwelle aufwiesen, keine Reduzierung der Höranstrengung nachgewiesen werden. Picou et al. [9] führten ein Dual-Task Paradigma mit Schwerhörigen, jeweils versorgt und unversorgt, durch. Die Primäraufgabe lag im Verstehen von Einsilbern, als Sekundäraufgabe wurde die Reaktionszeit auf einen visuellen Reiz gemessen. Der SNR der akustischen Kondition wurde so gewählt, dass 60% der Wörter richtig verstanden wurden. Eine Erhöhung der Reaktionszeit wurde in dieser Studie als Indiz für eine erhöhte Höranstrengung interpretiert. Unabhängig vom Versorgungsstatus konnte jedoch keine Abnahme der Höranstrengung bei audiovisueller Darbietung beobachtet werden. Auch Gosselin und Gagné [5] nutzten ein Dual-Task Paradigma zur Ermittlung der Höranstrengung, jedoch wurden ausschließlich Normalhörende getestet. Als Primäraufgabe wurde Wortverstehen und als Sekundäraufgabe das Erkennen taktiler Muster getestet. Der SNR wurde so gewählt, dass innerhalb der jeweiligen Kondition im Mittel ein Sprachverstehen von 80% erreicht wurde. Zur Auswertung wurden für die akustische sowie für die audiovisuelle Kondition jeweils die Kosten für die gleichzeitige Ausführung von zwei Aufgaben im Vergleich zur Ausführung einer der beiden Aufgaben einzeln berechnet. Aufgrund der erhöhten Kosten bei der audiovisuellen Kondition wurde diese Kondition als anstrengender gedeutet. Bei beiden Probandengruppen resultierte die Verwendung des Mundbildes in einer Zunahme der Höranstrengung, wobei die Zunahme bei den Älteren signifikant höher war. Aus diesen Studien ist zu schließen, dass besonders ältere Probanden, welche mitunter eine Schwerhörigkeit aufweisen, nicht immer von der audiovisuellen Darbietung profitieren. Dies ist in Einklang mit dieser Studie. In Kapitel „Vergleich der Konditionen“ und „Vergleich der Probandengruppen“ zeigte sich, dass die schwerhörigen Probanden, welche älter waren als die Normalhörenden, einen geringeren Vorteil durch die visuelle Information aufwiesen. Für die Schwerhörigen wären höhere SNR-Werte nötig, um im gleichen Ausmaß wie die Normalhörenden von der visuellen Darbietung zu profitieren.

Intra- und interindividuelle Standardabweichung

Die interindividuelle Standardabweichung war in der akustischen Kondition ca. dreimal höher als die intraindi-

viduelle Standardabweichung der gleichen Kondition. Bei der audiovisuellen Kondition unterschieden sich die Abweichungen etwa um den Faktor 4. Die interindividuelle Standardabweichung in der akustischen Kondition stimmt mit den Werten von Krueger et al. [15] überein. Obwohl die Probanden die gleiche Instruktion erhielten, ergaben sich zwischen den Probanden erkennbar unterschiedliche SNR-Werte für die Höranstrengungskategorien. Als Ursache kann davon ausgegangen werden, dass sowohl die Aufgabenstellung als auch die Skala mit den Worten „müheless“ und „extrem anstrengend“ von den Probanden unterschiedlich interpretiert wurden. Außerdem wurde zur Bewertung der Höranstrengung die Frage „wie anstrengend es [...] war, der Sprecherin zu folgen“ gestellt. Durch die Abfrage der subjektiven Höranstrengung ist nicht auszuschließen, dass Wahrnehmungen wie Lautheit oder Sprachverstehen indirekt mit in die Bewertung eingeflossen sind. Ein weiterer Einflussfaktor könnte sein, dass nicht jeder Proband den OLSA kannte. Aufgrund des Trainingseffektes des OLSA [27] kann vermutet werden, dass die Probanden, welche mit dem OLSA vertraut sind, anders bewerten als Probanden, denen der OLSA unbekannt ist. Möglicherweise reichte das durchgeführte Training in diesem Fall nicht aus. Außerdem waren die Standardabweichungen der akustischen Kondition tendenziell niedriger als in der audiovisuellen Kondition. Ein Proband berichtete, dass die audiovisuelle Kondition nicht sehr realistisch sei, da normalerweise der Pegel des Sprechers abnimmt, wenn der Abstand zum Zuhörer zunimmt. Der Pegel der Sprecherin wurde zwar verändert, der Abstand zum Zuhörer war jedoch stets derselbe. Es ist nicht auszuschließen, dass andere Probanden dies auch als störend empfunden haben.

Die intraindividuelle Standardabweichung lag in der akustischen Kondition zwischen 0,8 dB und 2,8 dB sowie in der audiovisuellen Kondition zwischen 1,7 dB und 3,4 dB und ist somit etwas geringer als bei Krueger et al. [15], welche in der akustischen Kondition Werte zwischen 1,0 dB und 3,8 dB erzielten.

Test-Retest Reliabilität

Unabhängig von der Kondition und den Maskierern wurde kein signifikanter Unterschied zwischen den Ergebnissen der beiden Messtermine ermittelt. Hingegen wurde bei Krueger et al. [15] ein signifikanter Unterschied in der Bewertung des IFFM zwischen dem ersten und zweiten Termin, nicht jedoch zwischen zweitem und drittem Termin, bestimmt. Der gemittelte IKK dieser Studie betrug 0,89 und ist damit mit Krueger et al. [15], welche einen Wert von 0,9 erzielten, vergleichbar. Werte zwischen 0,75 und 0,90 deuten auf eine gute Test-Retest Reliabilität hin [28].

Vergleich der Probandengruppen

Für die Schwerhörigen waren für die Kategorien 7 ESCU und 13 ESCU signifikant höhere SNR-Werte als bei den Normalhörenden nötig, um die gleiche Höranstrengungs-

bewertung zu erzielen. Bei der Kategorie 1 ESCU waren keine signifikanten Unterschiede nachweisbar. Ein Vergleich mit den Ergebnissen der akustischen Kondition von Krueger et al. [1] liefert beim Olnoise und IFFM bei 1 ESCU ebenfalls einen Unterschied von ca. 3 dB. Bei 13 ESCU weichen die Ergebnisse von Krueger et al. [1] ab. Die Differenz der Probandengruppe betrug dort beim IFFM 5 dB, während in dieser Untersuchung eine Differenz von fast 17 dB erreicht wurde. Während hier nur drei der neun Probanden mit Hörgeräten versorgt waren, handelte es sich bei Krueger et al. [1] ausnahmslos um erfahrene Hörgeräte-Träger. Es ist zu vermuten, dass der gewohnte Versorgungsstatus Einfluss auf die Messungen ohne Hörhilfen gehabt hat. Da Krueger et al. [1] einen Zusammenhang zwischen PTA4 und den Höranstrengungskategorien ermittelten, ist davon auszugehen, dass auch in dieser Untersuchung der Hörverlust Einfluss auf die Höranstrengungsbewertung hatte. Auch andere Studien belegten eine Abhängigkeit der Höranstrengung von einer Schwerhörigkeit [3], [4], [29]. Bernstein und Auer [30] zeigten, dass Schwerhörende, bei denen die Hörbeeinträchtigung bereits früh eingesetzt hatte, signifikant bessere Ergebnisse beim Sprachverstehen bei rein visueller Präsentation als Normalhörende erzielten. Aufgrund des Zusammenhangs von Sprachverstehen und Höranstrengung liegt die Vermutung nahe, dass diese Personen auch eine geringere Höranstrengung aufweisen. Da sowohl in den genannten Studien sowie in dieser Arbeit Schwerhörende weniger Nutzen durch das Absehen in Bezug auf die Höranstrengung aufwiesen, ist davon auszugehen, dass neben dem Hörvermögen auch Faktoren wie das Alter und kognitive Prozesse einen Einfluss auf die Höranstrengung haben.

Messverfahren ACALES

Das Messverfahren ACALES war für die Probanden leicht verständlich und einfach durchführbar. Die Probanden haben das Verfahren gut angenommen, sodass die Messungen ohne Auffälligkeiten durchgeführt werden konnten. Die Ermittlung einer Höranstrengungskurve in einer Kondition dauerte jedoch ca. zwischen zehn und zwölf Minuten, sodass ein Termin mit sechs Messungen bei einigen Probanden zur Ermüdung führte. Dennoch kann gesagt werden, dass ACALES aufgrund der guten Test-Retest-Reliabilität sowie der Möglichkeit zur Erfassung von individuellen Unterschieden sowohl in der akustischen als auch in der audiovisuellen Kondition für die Ermittlung subjektiver Höranstrengung eingesetzt werden kann. Vergleichende Messungen, z.B. über verschiedene Bedingungen, Versorgungen, oder verschiedene Tage hinweg, sind somit mit diesem Verfahren sehr gut möglich.

Anmerkungen

Interessenkonflikte

Die Autoren erklären, dass sie keine Interessenkonflikte in Zusammenhang mit diesem Artikel haben.

Danksagung

Die Untersuchung wurde im Rahmen des Projektes VIB-Hear durchgeführt, welches mit Mitteln des Europäischen Fonds für regionale Entwicklung (EFRE) und Mitteln des Landes Niedersachsen gefördert wird. Besonderer Dank gilt der Arbeitsgruppe von Volker Hohmann für die Bereitstellung des Video- und Audiomaterials.

Hinweis

Inhalte dieses Beitrags wurden bei der 22. DGA-Jahrestagung vom 6.–9. März 2019 in Heidelberg präsentiert.

Literatur

1. Krueger M, Schulte M, Zokoll MA, Wagener KC, Meis M, Brand T, et al. Relation Between Listening Effort and Speech Intelligibility in Noise. *Am J Audiol.* 2017;26(3S):378. DOI: 10.1044/2017_AJA-16-0136
2. McGarrigle R, Munro KJ, Dawes P, Stewart AJ, Moore DR, Barry JG, et al. Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group 'white paper.' *Int J Audiol.* 2014;53(7):433-45. DOI: 10.3109/14992027.2014.890296
3. Luts H, Eneman K, Wouters J, Schulte M, Vormann M, Buechler M, et al. Multicenter evaluation of signal enhancement algorithms for hearing aids. *J Acoust Soc Am.* 2010;127(3):1491-1505. DOI: 10.1121/1.3299168
4. Ohlenforst B, Zekveld AA, Lunner T, Wendt D, Naylor G, Wang Y, et al. Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation. *Hear Res.* 2017;351:68-79. DOI: 10.1016/j.heares.2017.05.012
5. Anderson Gosselin P, Gagné JP. Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. *Int J Audiol.* 2011;50(11):786-792. DOI: 10.3109/14992027.2011.599870
6. Anderson Gosselin P, Gagné JP. Older Adults Expend More Listening Effort Than Young Adults Recognizing Speech in Noise. *J Speech Lang Hear Res.* 2011;54(3):944. DOI: 10.1044/1092-4388(2010/10-0069)
7. Degeest S, Keppler H, Corthals P. The Effect of Age on Listening Effort. *J Speech Lang Hear Res.* 2015;58(5):1592-1600. DOI: 10.1044/2015_JSLHR-H-14-0288
8. Sarampalis A, Kalluri S, Edwards B, Hafter E. Objective Measures of Listening Effort: Effects of Background Noise and Noise Reduction. *J Speech Lang Hear Res.* 2009;52(5):1230. DOI: 10.1044/1092-4388(2009/08-0111)
9. Picou EM, Ricketts TA, Hornsby BWY. How Hearing Aids, Background Noise, and Visual Cues Influence Objective Listening Effort. *Ear Hear.* 2013;34(5):e52-e64. DOI: 10.1097/AUD.0b013e31827f0431

10. Volz L. Der Unterricht im Absehn der Sprache. Arch Für Ohren-Nasen- Kehlkopfheilkd. 1954;165(2-6):362-372. DOI: 10.1007/BF02134821
11. Leonhardt A. Einführung in die Hörgeschädigtenpädagogik: mit 88 Übungsaufgaben und zahlreichen Tabellen. 3., überarbeitete und erweiterte Auflage. München Basel: Reinhardt; 2010.
12. Bernstein JGW, Grant KW. Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. J Acoust Soc Am. 2009;125(5):3358-3372. DOI: 10.1121/1.3110132
13. Nirme J, Haake M, Lyberg Åhlander V, Brännström J, Sahlén B. A virtual speaker in noisy classroom conditions: supporting or disrupting children's listening comprehension? Logoped Phoniater Vocol. 2018;44(2):79-86. DOI: 10.1080/14015439.2018.1455894
14. Sommers MS, Phelps D. Listening Effort in Younger and Older Adults: A Comparison of Auditory-Only and Auditory-Visual Presentations. Ear Hear. 2016;37:62S-68S. DOI: 10.1097/AUD.0000000000000322
15. Krueger M, Schulte M, Brand T, Holube I. Development of an adaptive scaling method for subjective listening effort. J Acoust Soc Am. 2017;141(6):4680-4693. DOI: 10.1121/1.4986938
16. Wagener K, Kühnel V, Kollmeier B. Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests. Z Audiol. 1999;38(1):4-15.
17. Festen JM, Plomp R. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. J Acoust Soc Am. 1990;88(4):1725-1736. DOI: 10.1121/1.400247
18. Brungart DS. Informational and energetic masking effects in the perception of two simultaneous talkers. J Acoust Soc Am. 2001;109(3):1101-1109. DOI: 10.1121/1.1345696
19. Wagener KC, Hochmuth S, Ahrlich M, Kollmeier B. Der weibliche Oldenburger Satztest. In: Deutsche Gesellschaft für Audiologie, ed. Abstracts der 17. Jahrestagung der Deutschen Gesellschaft für Audiologie; 2014. Abrufbar unter/Available from: https://www.dga-ev.com/fileadmin/daten/downloads/bisherige_Jahrestagung/dga2014_programm_final.pdf
20. Llorach G, Kirschner F, Grimm G, Zokoll MA, Wagener KC, Hohmann V. Development and Evaluation of Video Recordings for the OLSA Matrix Sentence Test. ArXiv E-Prints. 2019; arXiv:1912.04700
21. Holube I, Fredelake S, Vlaming M, Kollmeier B. Development and analysis of an International Speech Test Signal (ISTS). Int J Audiol. 2010;49(12):891-903. DOI: 10.3109/14992027.2010.506889
22. Holube I. Speech intelligibility in fluctuating maskers. Proceedings of the International Symposium on Auditory and Audiological Research. 2011;(3):57-64.
23. Oetting D, Brand T, Ewert SD. Optimized loudness-function estimation for categorical loudness scaling data. Hear Res. 2014;316:16-27. DOI: 10.1016/j.heares.2014.07.003
24. Koelwijn T, Zekveld AA, Festen JM, Kramer SE. Pupil Dilation Uncovers Extra Listening Effort in the Presence of a Single-Talker Masker. Ear Hear. 2012;33(2):291-300. DOI: 10.1097/AUD.0b013e3182310019
25. Rennies J, Schepker H, Holube I, Kollmeier B. Listening effort and speech intelligibility in listening situations affected by noise and reverberation. J Acoust Soc Am. 2014;136(5):2642-2653. DOI: 10.1121/1.4897398
26. Schepker H, Haeder K, Rennies J, Holube I. Perceived listening effort and speech intelligibility in reverberation and noise for hearing-impaired listeners. Int J Audiol. 2016 Dec;55(12):738-747. DOI: 10.1080/14992027.2016.1219774
27. Schlueter A, Lemke U, Kollmeier B, Holube I. Normal and Time-Compressed Speech: How Does Learning Affect Speech Recognition Thresholds in Noise?. Trends Hear. 2016;20:1-13. DOI: 10.1177/2331216516669889
28. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-163. DOI: 10.1016/j.jcm.2016.02.012
29. Mackersie CL, MacPhee IX, Heldt EW. Effects of Hearing Loss on Heart Rate Variability and Skin Conductance Measured During Sentence Recognition in Noise. Ear Hear. 2015;36(1):145-154. DOI: 10.1097/AUD.0000000000000091
30. Auer ET, Bernstein LE. Enhanced Visual Speech Perception in Individuals With Early-Onset Hearing Impairment. J Speech Lang Hear Res. 2007 Oct;50(5):1157-1165. DOI: 10.1044/1092-4388(2007/080)

Korrespondenzadresse:

Saskia Ibelings
 Hörzentrum Oldenburg GmbH, Marie-Curie-Straße 2,
 26129 Oldenburg, Deutschland, Telefon: +49 441
 2172-100
 saskia.ibelings@uni-oldenburg.de

Bitte zitieren als

Ibelings S, Schulte M, Krüger M, Holube I. Audiovisuelle Realisierung des subjektiven Höranstrengungsmessverfahrens ACALES. GMS Z Audiol (Audiol Acoust). 2020;2:Doc04. DOI: 10.3205/zaud000008, URN: urn:nbn:de:0183-zaud0000083

Artikel online frei zugänglich unter

<https://www.egms.de/en/journals/zaud/2020-2/zaud000008.shtml>

Veröffentlicht: 24.04.2020

Copyright

©2020 Ibelings et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.

Audiovisual realization of the subjective listening effort measurement method ACALES

Abstract

Speech intelligibility can be positively influenced by the combination of auditory and visual speech information. Not only do people with hearing difficulties benefit from the additional information provided by the mouth image, but also those with normal hearing. However, the results regarding listening effort have so far been ambiguous. Although one study showed a decrease in listening effort, another showed an increase in listening effort when the mouth image was also presented. The aim of the current study was to measure subjective listening effort in an audiovisual situation when compared to a purely acoustic presentation. The adaptive scaling method ACALES (Adaptive CAtegorical Listening Effort Scaling, see Krueger et al., *J Acoust Soc Am.* 2017 06;141(6):4680–93) was used. Since ACALES has previously only been used for purely acoustic stimuli, it was first necessary to extend the method to play videos. In the acoustic as well as in the audiovisual condition, sentences of the Oldenburg sentence test (OLSA) were presented in the presence of different background noises. Additionally, in the audiovisual condition the corresponding mouth image of the speaker was shown on a screen. The measurements were performed with 15 young participants that had normal hearing and ten older participants with hearing impairment. Besides measuring listening effort in acoustic and audiovisual conditions for different maskers, the intra- and inter-individual standard deviation, as well as the test-retest-reliability, were determined. A dependence on the masker was found for both conditions. In addition to a significant difference between audiovisual and acoustic conditions, there was also a significant difference between the subject groups. The measurement method is suitable for recording interindividual differences in the evaluation of listening effort and has a good reliability

Keywords: listening effort, ACLES, audiovisual, hearing impairment, speech in noise

Introduction

and intelligible to the listener, the situation can still be perceived as tiring and strenuous. This is evident for both normal-hearing and hearing-impaired people, especially in situations where background noise is high or reverberation is strong [1]. As the signal-to-noise ratio (SNR) decreases, speech intelligibility decreases. At the same time, the concentration and effort required for comprehension increases steadily [1]. This additional effort can be described as listening effort [2]. In addition to background noise (masking) and the resulting SNR, listening effort can also be negatively influenced by factors such as hearing loss [1], [3], [4] and age [5], [6], [7]. The use of hearing aids can have a positive effect on listening effort [1], [3], [8], [9]. Another important, influential factor is the use of the mouth image. The additional information obtained through the mouth image can be used to com-

plete what was not understood and correct what was misunderstood [10], [11]. Studies have shown that speech intelligibility can be improved by using the mouth image [9], [12], [13]. However, the influence on listening effort is not clear. In contradiction to the finding of a decrease in listening effort [14], the use of the mouth image can also lead to an increase in listening effort [5] or no change [9].

The results of the studies mentioned above were produced using objective measurement methods (such as dual-task paradigms). In contrast, in this study an adaptive measurement procedure was used to evaluate the subjective listening effort (Adaptive CAtegorical Listening Effort Scaling, ACALES, [15]). In ACALES, sentences from the Oldenburg sentence test (OLSA) [16] are presented in background noise. The task of the participants is to evaluate the perceived listening effort on a 13-level category scale from “no effort” to “extreme effort” and a 14th additional category “only noise”. Each category is scored from 1 ESCU (“no effort”, unit: effort scaling categorical

Saskia Ibelings^{1,2}

Michael Schulte^{1,3}

Melanie Krüger^{1,3}

Inga Holube^{2,3}

1 Hörzentrum Oldenburg GmbH, Oldenburg, Germany

2 Institute of Hearing Technology und Audiology, Jade University of Applied Sciences, Oldenburg, Germany

3 Cluster of excellence “Hearing4All”, Oldenburg, Germany

unit) to 13 ESCU (“extreme effort”) or 14 ESCU (“only noise”). During the measurement, the SNR is adaptively changed depending on the response of the participants. The result is a function in which each listening effort category is assigned an SNR value [15].

In order to investigate the influence of the mouth image on subjective listening effort, it was necessary to extend the ACALES measurement method, which had previously only been used for purely acoustic stimuli, to incorporate the reproduction of audiovisual stimuli. Based on the findings of previous studies, it was assumed that the mouth image had an influence on listening effort [5], [14] and that there was a difference in listening effort scores between normal-hearing and hearing-impaired persons [1], [3], [4]. Furthermore, it was expected that temporally stationary maskers increase the listening effort more than temporally fluctuating ones [1], [15]. With fluctuating maskers, it is possible to listen in the temporal gaps, which improves speech intelligibility [17] and therefore probably reduces the listening effort. Furthermore, since maskers from a source having the same sex as the OLSA speaker have a similar frequency spectrum, it was assumed that these maskers lead to a greater listening effort than maskers with the spectrum of a speaker of the opposite sex [18]. To evaluate the extended measurement procedure, the intra- and inter-individual standard deviations and test-retest reliability were determined. The study thus had the following purposes:

- Determination of the subjective listening effort
 - for purely acoustic and audiovisual stimuli
 - for different maskers
 - for normal-hearing and hearing-impaired people
- Determination of the intra- and inter-individual standard deviation
- Determination of the test-retest reliability

The investigation was limited to the analysis of listening effort. Although the relationship between listening effort and speech intelligibility would have been of interest under these conditions, the simultaneous measurement of speech intelligibility was not performed because of time and effort required. Since it is known that visual speech features can significantly improve speech intelligibility, which in turn influences listening effort, this circumstance was taken into account through appropriate references at several points in this paper.

Material and method

Participants

A total of 25 participants took part in the measurements. The first group included 15 normal-hearing (NH) participants aged 19 to 27 years (22.7 ± 2.3 years). Seven of these were male and eight female. As used by Krueger et al. [15], normal hearing was defined by a PTA_4 (averaged air conduction thresholds at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz) of better than 20 dB HL.

The averaged PTA_4 was 2.0 ± 2.8 dB HL in the right ear and 2.3 ± 3.2 dB HL in the left ear. Two participants already had experience with the ACALES measurement procedure. One subject did not know the OLSA stimuli.

Ten hearing-impaired persons (HI) aged 63 to 76 years (70.1 ± 4.2 years) represented the second group of participants. Six of the participants were male, four female. Their average PTA_4 was 43.1 ± 5.3 dB HL in the right ear and 40.3 ± 5.3 dB HL in the left ear. Three of the participants had been fitted with hearing aids, but all measurements were performed without these. Most of the participants were familiar with the OLSA, but none of them had any experience in measuring listening effort.

Stimuli

Sentences from OLSA with a female speaker [19] were used as stimuli in both the acoustic and audiovisual conditions. The sentences of the OLSA use the same order of word categories (name-verb-number-adjective-objective), e.g. “Nina paints ten wet armchairs” (in German: “Nina malt zehn nasse Sessel”). Each word category consists of ten different words [16]. In the audiovisual condition, the corresponding mouth movements of the speaker were simultaneously shown on a screen. The speaker, who had also recorded the female OLSA, was filmed afterwards [20]. She was visible from the shoulder up, in front of a green screen as background. In both conditions, different maskers were additionally presented.

Maskers

The measurements were conducted using three different maskers. Besides the female Olnoise (Olnoise with the spectrum of a female speaker [16], [19]), the International Female Fluctuation Masker (IFFM; [22]) and OLSA sentences (male) [16] were used as an interfering speaker (interferer). The Olnoise is a stationary noise generated by multiple random overlays of OLSA sentences [18]. Thus the averaged long-term spectrum of the Olnoise is identical to that of the stimuli. The IFFM is a variant of the International Speech Test Signal (ISTS, [21]). The pauses were shortened to 250 ms [22]. For the ISTS and IFFM recordings of six female speakers of different native languages were used (American English, Arabic, Mandarin, German, French, and Spanish). Due to the segmentation and mixing of the signals, the masker is mostly unintelligible. The averaged long-term spectrum is equivalent to that of a female speaker [21]. For the interfering speaker, 15 sentences of OLSA from a male speaker [16] were concatenated. The average time interval between the sentences was about 500 ms. Consequently, there was a time gap between the sentences without spoken words.

Equipment

The measurements took place in a soundproof room. The D/A converted signals (sound card ADI-8 PRO by RME; Haimhausen, Munich, Germany) were presented via a loudspeaker (Mackie HR 824; Bothell, Washington, USA) that was placed about 1.3 m in front of the participants. In addition, a touch screen was in front of the participants, on which the scale for the rating of the listening effort was shown. This screen was also used in the audiovisual condition to present the videos. Both the acoustic and the audiovisual stimuli were presented via the VLC media player (VideoLan, version 3.0.3; Verden, Germany). The synchronization of audio and video signal was performed by eye. Listening effort measurements were performed using ACALES [15], which was implemented in MATLAB (version 2007b; Natick, Massachusetts, USA).

Listening effort rating

ACALES [15] was used to measure the subjective listening effort. For each SNR, three random sentences of the OLSA were presented. The level of the target sentences was changed adaptively according to the participants' responses, whereas the level of the masker was constant at 65 dB SPL. In order to avoid too-loud or too-soft levels, the SNR range was limited to -35 to 25 dB SPL. For the normal-hearing participants, the starting SNR was set at 0 dB, but at 10 dB for the hearing-impaired participants. The subjective listening effort was evaluated on a 14-point scale from "no effort" (1 ESCU) to "only noise" (14 ESCU). In the first phase of ACALES, the SNR values were determined for the limits "no effort" and "extreme effort". In the second phase, seven different SNR values for the named categories were presented in random order within the previously determined SNR range. After a recalculation of the limits, SNR values for the six unnamed intermediate categories were randomly presented twice in the third phase. In total, there were at least 21 SNR presentations per measurement. Figure 1 shows the scales employed.

Since in the audiovisual measurement condition it could not be excluded that the participants were able to follow the mouth image of the speaker, although the signal was acoustically not perceptible, the category "only noise" was renamed to "nothing perceptible". This category means that the sentences or certain segments of the sentences were not accessible acoustically or visually via the mouth image. The meanings of the categories "noise only" and "nothing perceptible" were explained in detail to the participants before the measurements.

Measurement procedure

The study was approved by the ethics committee in accordance with the ethics application Drs. 47/2017. The measurement procedure is shown schematically in Figure 2. The air-conduction hearing threshold was measured after receiving information and written consent

of the participants. Subsequently, the subjective listening effort was measured. The measurements of the hearing-impaired participants were performed without hearing aids. In one session, the listening effort in both conditions (acoustic and audiovisual) was determined for all maskers. In addition to the order of the conditions, the order of the maskers within a condition was randomized. In the first measurement condition, each of the maskers was trained before the actual measurement in order to become familiar with the stimuli, maskers and the measurement procedure. In the second condition, the maskers, measurement procedures and OLSA sentences were already known and only one training session was performed at the start to familiarize participants with the new condition (audiovisual/acoustic). The training was similar to the first measurement phase of ACALES. The stimuli were presented in random order both during training and during the measurement. For ten of the 15 normal-hearing participants, the listening effort in both conditions was again evaluated in a second session and using all maskers. Only one training session per condition was performed during this session, as the measurement procedures, stimuli and masker were assumed to be known. On the first day, four training sessions and six measurements were performed per subject. On the second day, the number of training sessions was reduced to two.

Analysis and statistics

For all participants, listening-effort curves were fitted with the BX fitting method integrated in ACALES [23]. Krueger et al. [1], [15] have already shown that this fitting method generally leads to a valid adaptation of the curves. These curves are linear from 1 ESCU to 7 ESCU and from 7 ESCU to 13 ESCU. The point of intersection of the lines was smoothed between 5 ESCU and 9 ESCU [15]. The listening-effort curves were averaged over all participants per condition (i.e. presentation mode), masker and listening-effort category.

The results were analyzed with SPSS 25.0.0. According to the Kolmogorov-Smirnov test, all data were normally distributed. One normal-hearing participant was excluded from the evaluation as an outlier because the data deviated by more than three times the standard deviation. One of the hearing-impaired participants also had to be excluded because the listening effort was never rated lower than 4 ESCU, independent of the condition and the maskers. Consequently, a valid fitting of the listening effort curve was not possible, since the listening effort of this subject was not represented by the values calculated with the BX-fit.

The statistical tests used were analysis of variance (ANOVA) for repeated measurements with a significance level of $\alpha=0.05$. As within-subject factors, the named categories (1 ESCU, 3 ESCU, 5 ESCU, 7 ESCU, 9 ESCU, 11 ESCU and 13 ESCU) and the maskers (Olnoise, IFFM, interferer) were always chosen. Depending on the research question, there was an additional within-subject

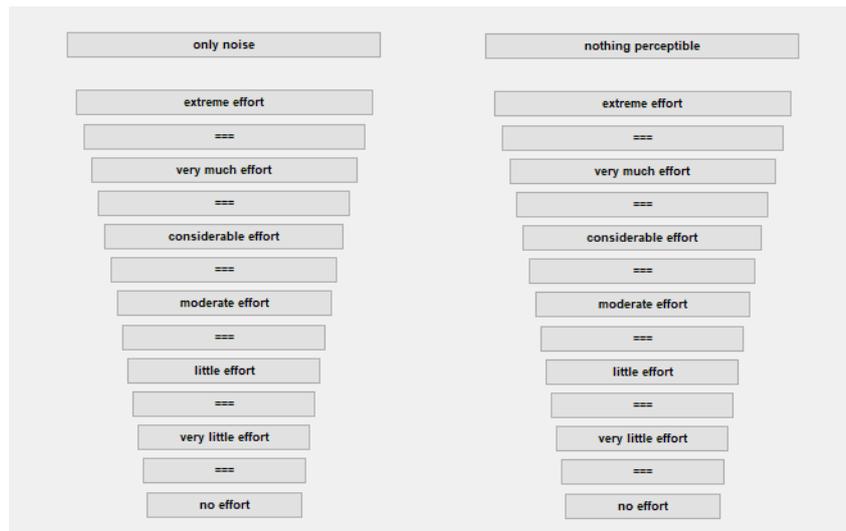


Figure 1: Scales to evaluate the subjective listening effort

Left: Scale for the acoustic condition, right: Scale for the audiovisual condition; the category “only noise” was renamed “nothing perceptible”.

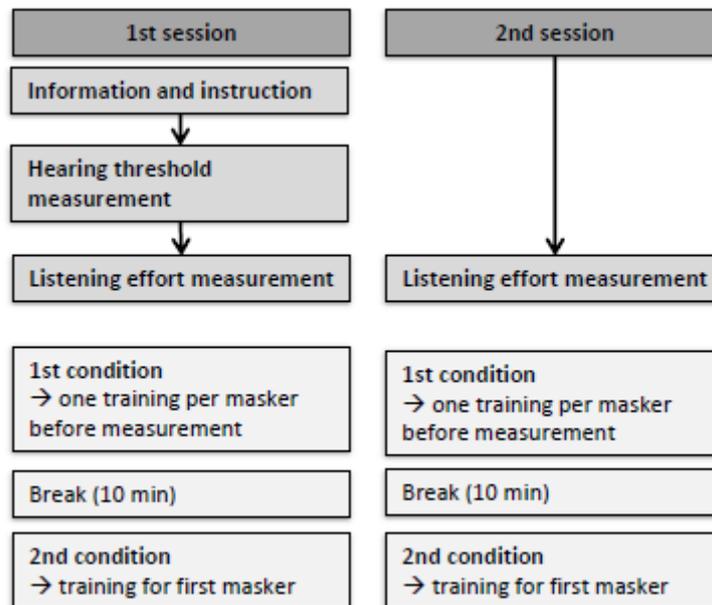


Figure 2: Schematic illustration of the measurement procedure

The orders of the conditions (acoustic and audiovisual) as well as the maskers (Olnoise, IFFM, interferer) were randomized. Ten of the 15 normal-hearing participants took part in the second session.

factor condition (acoustic, audiovisual) or time (1st session, 2nd session), or a between-subject factor participant group (normal hearing, hearing-impaired). Based on the linear regressions of the listening effort functions, post-hoc t-tests were performed for the categories 1 ESCU, 7 ESCU and 13 ESCU. Nine t-tests per condition were performed based on the three selected categories and the three maskers. For the same reason, nine t-tests were also performed for each condition, to compare the conditions and the results of the first and second sessions. Therefore, the significance level after a Bonferroni correction was in all cases $\alpha=0.05/9=0,00\bar{5}$.

Results

Listening effort in the acoustic condition

In the normal-hearing participants, a dependence of the rated listening effort on the masker was seen in the acoustic (a) condition (see Figure 3, above). A shift of the curves to lower SNR values means less listening effort. At low SNR values, the Olnoise was perceived as most effortful, while the masker was perceived as least effortful. Repeated measures ANOVA confirmed the difference in the ratings of the maskers [Greenhouse Geisser $\epsilon=0.682$, $F(1.364, 17.735)=21.481$, $p<0.001$]. Furthermore, a significant difference between the categories

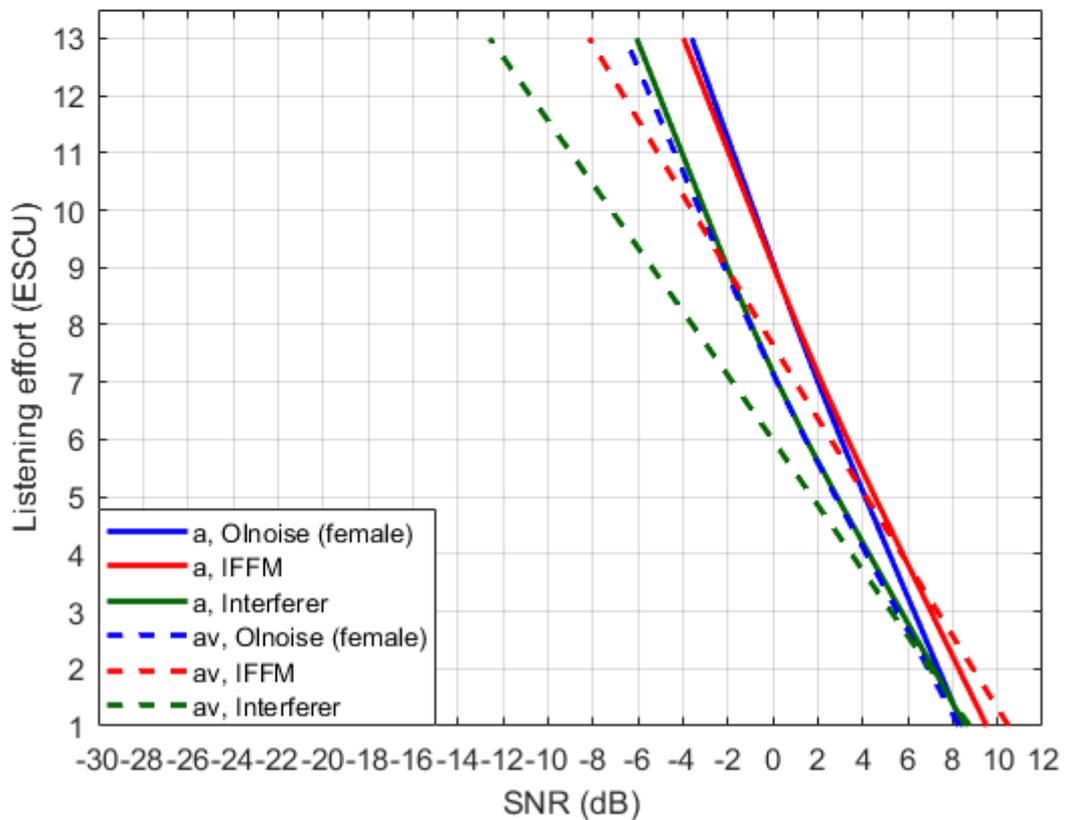
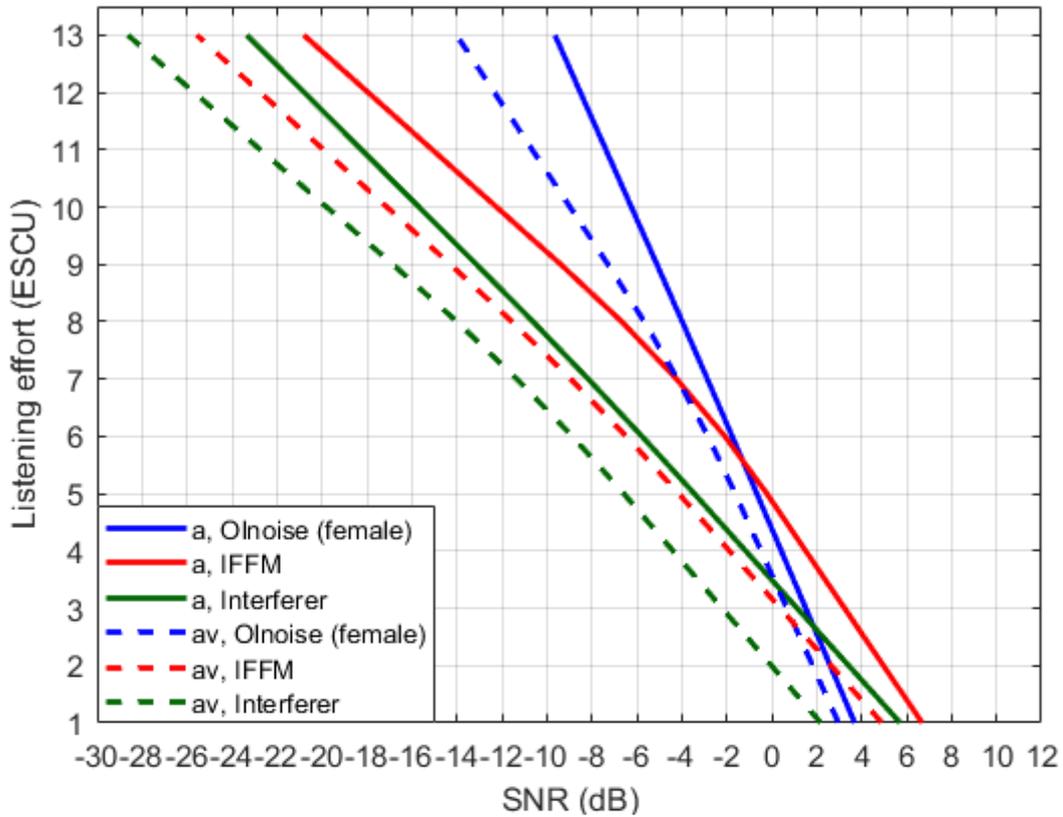


Figure 3: Top: averaged listening effort curves of normal-hearing persons (n=14); bottom: averaged listening curves of the hearing impaired (n=9)

The solid lines represent the results of the acoustic (a) condition and the dotted lines the results of the audiovisual (av) condition.

[$F(6, 78)=380.456, p<0.001$] was observed. An interaction between masker and categories was also found [$F(12, 156)=77.175, p<0.001$]. Using the Bonferroni correction, t-tests for paired samples between Olnoise and IFFM showed a significant difference for the category “extreme effort”. Olnoise and interferer also showed a significant difference in the category “moderate effort” (7 ESCU). At 7 ESCU, IFFM and interferer also differed significantly (all $p<0.001$). In the category “no effort” (1 ESCU), there was a significant difference between Olnoise and IFFM ($p=0.018$).

In contrast, the listening-effort curves of the maskers showed little difference in the acoustic condition for the hearing-impaired participants (see Figure 3, below). Repeated measures ANOVA showed a significant difference between the categories [$F(2.6)=3.292, p<0.001$], but no significant difference was found in the masker ratings [$F(6.48)=143.908, p=0.063$]. A significant interaction between masker and category was also observed [$F(12.96)=2.710, p=0.003$]. This indicates that the differences in the rating of the maskers depended on the listening-effort category. Differences can be seen at low SNRs, i.e. high listening-effort categories. The lower the SNR range, the higher the SNR differences between the maskers needed to achieve the same listening-effort rating.

Listening effort in the audiovisual condition

In the audiovisual (av) condition, a different rating of the maskers could be seen among normal-hearing participants (see Figure 3, above). Overall, the lowest SNR values were achieved by the interferer, so that it was perceived as the least effortful, while the Olnoise was perceived as the most effortful, due to having the highest SNR values. The differences between the maskers increased with increasing listening effort. Between Olnoise and interferer, the difference in SNR at 13 ESCU was about 14 dB. For the IFFM, SNR values that were on average 2.5 dB higher than those of the interfering speaker were required.

A repeated-measures ANOVA confirmed a significantly different rating of the maskers [$F(2, 26)=15.774, p<0.001$], as well as a significant difference in the categories [$F(6, 78)=334.995, p<0.001$]. The interaction between masker and categories was also significant [$F(12, 156)=25.674, p<0.001$]. The t-tests for paired samples showed significant differences between Olnoise and IFFM at 7 ESCU ($p=0.004$) and 13 ESCU ($p<0.001$), and between Olnoise and interferer, also in the categories 7 ESCU and 13 ESCU ($p<0.001$ each), considering the Bonferroni correction. For IFFM and interfering speaker, however, no significant differences were found ($p>0.05$). Among the hearing-impaired participants also, the interferer tended to be rated as least effortful (see Figure 3, below). In the comparison between IFFM and interferer, the difference up to 7 ESCU was approx. 2 dB, and up to 13 ESCU the difference increased to about 4 dB. No clear

difference was visible between Olnoise and IFFM. The ANOVA showed that the ratings of the maskers [$F(2.16)=6.393, p=0.009$] and the categories [$F(6.48)=123.614, p<0.001$] differed significantly. Furthermore, a significant interaction between category and masker was found [$F(12.96)=8.640, p<0.001$]. Post-hoc t-tests with a Bonferroni-corrected significance level showed that the SNR values of the interferer and Olnoise at 13 ESCU ($p=0.002$) and the interferer and IFFM at 7 ESCU ($p=0.004$) differed significantly.

Comparison of the conditions

For both normal-hearing and hearing-impaired participants, a shift in listening effort curves of the audiovisual condition relative to the acoustic condition towards lower SNR values was seen (see Figure 3). The audiovisual condition was thus perceived as less effortful than the acoustic condition. For normal-hearing participants, the curves of the conditions for IFFM and interferer are almost parallel. To achieve the same listening effort rating for the acoustic condition, these maskers required SNR values that were, on average, approx. 4 dB higher. For the Olnoise, the difference was about 1 dB up to 5 ESCU and increased to 4 dB at 13 ESCU. The repeated measures ANOVA confirmed a significant difference in the rating of the conditions [$F(1, 13)=14.656, p=0.002$]. In addition, significant interactions were found between category and masker [$F(12, 156)=64.328, p<0.001$] and condition and category [$F(6.78)=3.630, p<0.003$]. The post-hoc t-tests for paired samples (see Table 1, considering Bonferroni correction) showed significant differences in conditions for the IFFM at 7 ESCU ($p=0.004$) and for the interferer at 13 ESCU ($p<0.001$). In contrast, no significant influence of the mouth image was seen for the Olnoise. For the hearing-impaired participants, independently of the masker, the difference in conditions increased with increasing listening effort. While the difference of the conditions at 13 ESCU was only 2 dB for the Olnoise, it reached 4 dB for the IFFM and 6 dB for the interferer. However, an ANOVA for repeated measurements found no significant influence of the conditions on the evaluation of the listening effort [$F(1.8)=4.760, p=0.061$]. However, significant interactions were found, not only between condition and category and masker and category ($p<0.001$ each), but also between condition, masker and category ($p=0.012$). Due to the significant interactions with the condition factor, t-tests were performed for paired samples. For all maskers, a significant effect on the additional visual presentation of the mouth image was found at 13 ESCU (each $p<0.01$ considering Bonferroni correction).

Comparison of the participant groups

Figure 4 shows the comparison of the groups of participants. The hearing-impaired participants tended to show similar listening effort ratings to those of normal-hearing participants, independent of condition and masker at

Table 1: Results of post-hoc tests comparing the acoustic (a) and the audiovisual (av) conditions of normal-hearing (NH) and hearing-impaired participants (SH)

	Olnoise		IFFM		Interferer	
	NH	HI	NH	HI	NH	HI
1 ESCU	0.306	0.876	0.109	0.636	0.007	0.898
7 ESCU	0.021	0.040	0.004*	0.469	0.038	0.149
13 ESCU	0.009	< 0.001*	0.051	0.003*	< 0.001*	0.002*

* Statistically significant using the Bonferroni correction ($\alpha=0.005$)

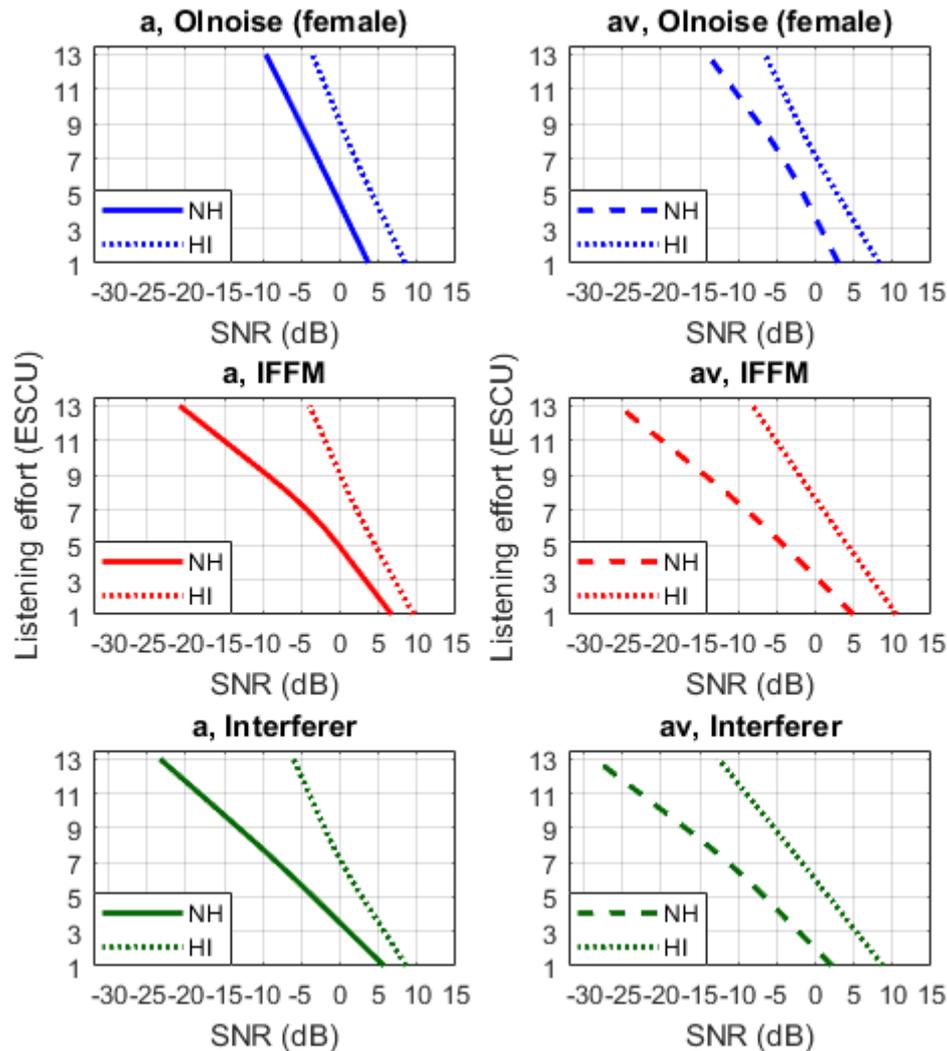


Figure 4: Comparison of the listening-effort curves of normal-hearing (n=14; dashed) and hearing-impaired participants (n=9; dotted) for all maskers in acoustic (a) and audiovisual (av) conditions

higher SNR values. While the curves for the Olnoise were shifted almost constantly by about 5 dB in both conditions, the difference for the other maskers and conditions increased with increasing listening effort up to about 17 dB each at 13 ESCU. The repeated measures ANOVA with the additional between-subject factor participant group (normal-hearing, hearing-impaired) confirmed a significant difference between the participant groups. Furthermore, there were significant interactions between masker and participant group, category and participant group, masker, category and participant group, as well as condition, masker, category and participant group (all $p < 0.05$). The t-tests for independent samples with a

Bonferroni-corrected significance level showed that the ratings of normal-hearing and hearing-impaired participants differed significantly at 7 ESCU and 13 ESCU, independent of condition and masker. A significant difference was also seen for the Olnoise in the acoustic condition at 1 ESCU (in each case $p < 0.05$ considering the Bonferroni correction).

Intra- und interindividual standard deviation

The intra- and inter-individual standard deviations were determined for the measurement results of ten normal-

Table 2: Intra- and interindividual standard deviation of the acoustic (a) and audiovisual (av) condition

		Intra-individual standard deviation in dB				Inter-individual standard deviation in dB			
		Olnoise	IFFM	Interferer	Mean	Olnoise	IFFM	Interferer	Mean
1 ESCU	a	2.2	1.4	1.9	1.8	3.4	4.2	5.3	4.3
	av	3.1	2.8	1.9	2.6	5.4	6.7	7.8	6.6
3 ESCU	a	1.4	1.6	2.0	1.8	3.3	4.6	5.5	4.5
	av	2.9	2.7	1.7	2.4	6.1	7.3	8.1	7.2
5 ESCU	a	1.5	1.9	2.4	1.9	3.2	5.0	5.8	4.7
	av	2.6	2.6	2.1	2.4	6.9	7.9	8.6	7.8
7 ESCU	a	1.3	2.1	2.5	2.0	3.1	5.4	6.0	4.8
	av	2.4	2.5	2.5	2.5	7.6	8.4	8.8	8.3
9 ESCU	a	1.0	2.2	2.4	1.9	2.9	5.6	6.0	4.8
	av	2.0	2.2	2.7	2.3	8.1	8.7	8.9	8.6
11 ESCU	a	0.8	2.4	2.3	1.9	2.8	5.7	5.9	4.8
	av	1.7	1.9	2.9	2.2	8.4	8.7	8.7	8.6
13 ESCU	a	1.0	2.8	2.3	2.9	2.8	5.9	5.9	4.9
	av	1.7	2.0	3.4	2.4	8.8	8.8	8.5	8.7

hearing participants. For this purpose, the SNR values of the named categories were used for each measurement condition and each masker. The intra-individual standard deviation was calculated from the SNR values of the two measurement sessions for each participant. The results for one listening effort category, one masker and one presentation condition were determined by averaging the intra-individual standard deviations of all participants. For the inter-individual standard deviation, the standard deviation of the SNR values of all participants per measurement session was determined for each condition, each masker and each named category and then averaged over the two measurement sessions. Table 2 contains the calculated values. For the intra-individual standard deviation, values were between 1.8 and 2.6 dB and the inter-individual standard deviation ranged from 4.3 to 8.7 dB. For both standard deviations, the values in the audiovisual condition tend to be higher than in the acoustic condition.

Test-retest reliability

Figure 5 shows the listening effort curves of the two measurement sessions per condition and masker. On average, almost identical results were obtained in the second session. To analyze the test-retest reliability, a repeated-measures ANOVA with the additional factor time (1st session, 2nd session) was performed. There was no significant difference between the results of the two measurement sessions [$F(1.9)=0.205, p=0.662$]. Furthermore, no significant interactions with the factor time could be confirmed (all $p>0.05$).

In addition, the intra-class correlation coefficient (ICC) was examined for each condition and each masker of the named categories (see Table 3). At 0.856, the averaged

ICC of the acoustic condition was slightly lower than that of the audiovisual condition (0.915). The lowest values were found in both the acoustic (0.699) and the audiovisual condition (0.717) for the Olnoise, both at 1 ESCU. The highest ICC was achieved for the audiovisual condition in Olnoise at 13 ESCU (0.983).

Discussion

Comparison of the maskers

The results show that the rating of listening effort in both acoustic and audiovisual conditions is dependent on the masker. However, significant differences were only found with increasing listening effort. The Olnoise was perceived to be the most effortful, whereas the interferer was perceived to be the least effortful. One explanation for the different perceptions lies in the temporal and spectral structure of the maskers. The long-term spectrum of the Olnoise is identical to that of the stimuli, so that the degree of masking is maximal [16]. Maskers with the same sex characteristics as the target signal also have a higher masking effect [18]. The IFFM also corresponds to female speech in the long-term spectrum, but fluctuates [22], so it is possible to listen into the temporal gaps of the masker [17]. The interfering speaker, however, is male, so that the masking effect is already reduced. In addition, the pauses of this masker are longer than the gaps in the IFFM. As a result, stimuli and masker are more easily distinguishable for the interfering speaker, so that the subjective listening effort is lowest. Overall, all the factors mentioned above can have an influence on speech intelligibility and, as a result, an influence on the listening effort.

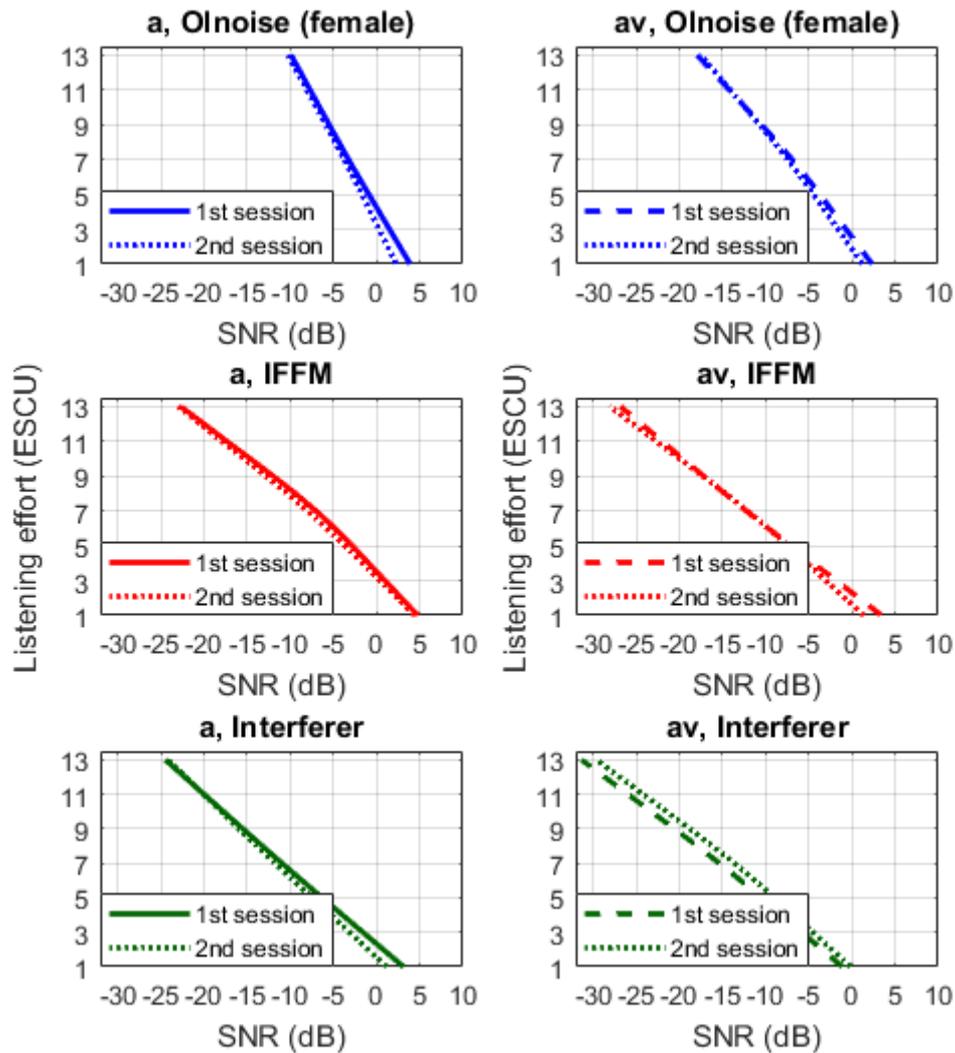


Figure 5: Comparison of the averaged listening-effort curves ($n=10$ normal-hearing participants) for all maskers in the acoustic (a) and in the audiovisual (av) conditions for the 1st and 2nd session

Table 3: Intra-class correlation coefficients for the acoustic (a) and the audiovisual (av) condition
Model: two way mixed, type: absolute agreement for average measures

		Intra-class correlation coefficients							Mean
		1 ESCU	3 ESCU	5 ESCU	7 ESCU	9 ESCU	11 ESCU	13 ESCU	
Olnoise	a	0.699*	0.747*	0.782*	0.806*	0.847*	0.892*	0.894*	0.810*
	av	0.717	0.851*	0.915*	0.948*	0.970*	0.982*	0.983*	0.909*
IFFM	a	0.893	0.921*	0.930*	0.924*	0.893*	0.848*	0.827*	0.891*
	av	0.821	0.849*	0.933*	0.951*	0.962*	0.966*	0.957*	0.920*
Interferer	a	0.885	0.906*	0.908*	0.901*	0.884*	0.836*	0.762*	0.869*
	av	0.941	0.959*	0.963*	0.956*	0.930*	0.874*	0.795*	0.917*

* Significant at $\alpha=0.05$

One exception to this are the hearing-impaired participants in the acoustic condition, where there was no significant difference in the rating of the listening effort between the maskers. One possible explanation is that the temporal and spectral resolution is limited or even impossible due to the hearing loss [17].

Krueger et al. [1], [15] also reported that stationary maskers (Olnoise) are perceived as more effortful than fluctuating ones (IFFM and ISTS). Even with comparable speech intelligibility, differences in the listening-effort rating of the maskers were detectable [1]. In another study, a higher listening effort was found when using an interfering speaker than with fluctuating or stationary

noise [24]. In that study by Koelewijn et al. [24] instead of a subjective method, pupillometry was used to determine the objective listening effort. In addition, the interfering speaker was spectrally modified, so that this masker had the same long-term spectrum as the target speaker. Besides the energetic masking, the informational masking is also present in the interfering speaker, which increases the masking effect and makes the interfering speaker more effortful. Fluctuating and stationary noise was therefore less effortful.

Comparison of the acoustic and audiovisual conditions

As expected, the listening effort rating of the acoustic and audiovisual conditions differed significantly in both groups of participants. The use of the mouth image reduced the subjective listening effort. However, the advantage of the mouth image only became apparent at negative SNR values, i.e. at higher listening effort. At higher SNR values, speech intelligibility is high [1], so that it can be assumed that the mouth image provides little additional information to improve speech intelligibility and to reduce listening effort. At lower SNR levels, the mouth image provides additional visual information to supplement and correct misunderstandings [11]. Additionally, visual speech can help to differentiate between target and interfering signals [12]. Both factors probably led to an increase in speech intelligibility. According to the results obtained here, the latter factor seems to have a greater effect on the interfering speaker than on the IFFM.

Llorach et al. [20], who used the same stimuli as in this study, found an improvement of the threshold for speech intelligibility of 80% by about 4.5 dB SNR when the mouth image was presented. Overall, it can be assumed that the improved speech intelligibility reduced the listening effort [1]. Furthermore, the significantly steeper curve for the O_{noise} compared to both the IFFM and the interferer indicates a relationship between speech intelligibility and listening effort. This is because the O_{noise} also has the steepest function when testing speech intelligibility [1]. For both normal-hearing and hearing-impaired participants, studies that have investigated the influence of noise and reverberation on speech intelligibility and subjective listening effort have also shown a decrease in listening effort with increasing speech intelligibility [25], [26].

Not only can subjective evaluation of listening effort be used, but also objective methods. Sommers and Phelps [14] investigated memory performance in young and older normal-hearing participants. They used word lists of varying lengths in quiet, where the last three words had to be repeated. The speech level was not reported, but it can be speculated that this was 60 dB SPL, as in the speech intelligibility measurement conducted by Sommers and Phelps [14]. An increase in correctly repeated words in the audiovisual condition indicated a decrease in listening effort. In contrast to the young participants, no reduction in listening effort could be detected in the older

participants, who had a significantly worse hearing threshold. Picou et al. [9] performed a dual-task paradigm with hearing-impaired participants, both with and without hearing aids. The primary task was to understand monosyllables; the secondary task was the measurement of response time to a visual stimulus. The SNR of the acoustic condition was chosen so that 60% of the words were correctly understood. In this study, an increase in reaction time was interpreted as an indication of increased listening effort. However, independent of the fitting status, no decrease in listening effort was found for audiovisual presentations. Gosselin and Gagné [5] also used a dual-task paradigm to measure listening effort, but only normal-hearing participants were tested. The primary task was word recognition and the secondary task was the recognition of tactile patterns. The SNR was selected to achieve an average speech comprehension of 80% within the respective condition. For the evaluation, the costs for the simultaneous performance of two tasks were calculated separately for both the acoustic and the audiovisual condition, compared to the performance of only one of the two tasks. Due to the increased costs for the audiovisual condition, this condition was interpreted as more effortful. In both groups of participants, the use of the mouth image resulted in an increase in listening effort, with a significantly higher increase in the elderly. From these studies, it can be concluded that especially older participants, who sometimes have a hearing loss, do not always benefit from an audiovisual presentation. This is in line with the current study. In the sections “Comparison of conditions” and “Comparison of participant groups” it was shown that hearing-impaired participants, who were older than normal-hearing ones, had less benefit from the visual information. For the hearing-impaired participants, higher SNR values would be necessary to benefit from the visual presentation to the same degree as the normal-hearing participants.

Intra- und inter-individual standard deviation

The inter-individual standard deviation in the acoustic condition was about three times larger than the intra-individual standard deviation for the same condition. In the audiovisual condition, the deviations differed by a factor of about 4, and the inter-individual standard deviation in the acoustic condition was in agreement with the values of Krueger et al. [15]. Although the participants received the same instructions, the SNR values for the listening-effort categories differed noticeably between the participants. The reason for this can be assumed to be that both the task and the scale with the words “no effort” and “extreme effort” were interpreted differently by the participants. Moreover, the question “How much effort does it require for you to follow the speaker?” was asked to evaluate the listening effort. By measuring the subjective listening effort, it cannot be excluded that perceptions such as loudness or speech intelligibility were indirectly included in the evaluation. Another influencing factor

could be that not every participant knew the OLSA. Due to the training effect of OLSA [27], it can be assumed that participants who are familiar with OLSA evaluate differently than participants who are not familiar with it. It is possible that in this case the training carried out was insufficient. Furthermore, the standard deviations of the acoustic condition tended to be lower than in the audiovisual condition. One participant reported that the audiovisual condition was not very realistic, because normally, the level of the speaker decreases as the distance to the listener increases. Although the level of the speaker was changed, the distance to the listener was always the same. It cannot be ruled out that other participants also perceived this as irritating.

The intra-individual standard deviation was between 0.8 dB and 2.8 dB in the acoustic condition, and between 1.7 dB and 3.4 dB in the audiovisual condition, and is thus somewhat lower than in Krueger et al. [15], who reported values between 1.0 dB and 3.8 dB in the acoustic condition.

Test-retest reliability

Independent of the condition and the maskers, no significant difference was found between the results of the two measurement sessions. However, Krueger et al. [15] noted a significant difference in the evaluation of IFFM between the first and second sessions, but not between the second and third sessions. The average ICC of this study was 0.89, and is thus comparable with Krueger et al. [15], which reported a value of 0.9. Values between 0.75 and 0.90 indicate a good test-retest reliability [28].

Comparison of the participant groups

For the hearing-impaired participants to achieve the same listening effort score, significantly higher SNR values were required for categories 7 ESCU and 13 ESCU than for normal-hearing participants. A comparison with the results of the acoustic condition of Krueger et al. [1] also shows a difference of about 3 dB for Olnoise and IFFM at 1 ESCU. At 13 ESCU, the results of Krueger et al. [1] differ. The difference of the group of participants in the IFFM was 5 dB, while in the current study, a difference of almost 17 dB was achieved. While here only three of the nine participants were fitted with hearing aids, the participants of the study Krueger et al. [1] were – without exception – experienced hearing-aid users. It can be assumed that the usual fitting status had an influence on the measurements without hearing aids. Since Krueger et al. [1] found a correlation between PTA_4 and the listening-effort categories, it can be assumed that hearing loss also had an influence on the listening-effort rating in the current study. Other studies have also shown a dependence of listening effort on hearing loss. [3], [4], [29]. Bernstein and Auer [30] reported that hearing-impaired people who started to suffer from hearing loss at an early age achieved significantly better results in speech intelligibility with purely visual presentation than those

of normal-hearing subjects. Given the relationship between speech intelligibility and listening effort, it is reasonable to assume that these persons also experience less listening effort. Since in both the studies mentioned above and in this paper hearing-impaired persons showed less benefit from visual presentation in terms of listening effort, it can also be assumed that factors such as age and cognitive processes have an influence on listening effort.

Measuring method ACALES

The ACALES measurement procedure was easy to understand and simple to carry out. Participants accepted the procedure well, so that the measurements could be performed without any abnormalities. However, the measurement of a listening effort curve in one condition took about ten to twelve minutes, so that a session with six measurements led to fatigue in some participants. Nevertheless, ACALES can be used for the evaluation of subjective listening effort based on its good test-retest reliability and the possibility of recording individual differences in both the acoustic and audiovisual conditions. Comparative measurements, e.g., across different conditions, fittings, or different days, are thus very possible with this method.

Notes

Competing interests

The authors declare that they have no competing interests.

Acknowledgement

The study was carried out as part of the VIBHear project, which is funded by the European Regional Development Fund (EFRE) and the State of Lower Saxony. Special thanks go to the working group of Volker Hohmann for providing the video and audio material. English language services were provided by <http://stels-ol.de/>.

Notification

The contents of this contribution were presented at the 22nd DGA Annual Conference from 6–9 March 2019 in Heidelberg.

References

1. Krueger M, Schulte M, Zokoll MA, Wagener KC, Meis M, Brand T, et al. Relation Between Listening Effort and Speech Intelligibility in Noise. *Am J Audiol*. 2017;26(3S):378. DOI: 10.1044/2017_AJA-16-0136
2. McGarrigle R, Munro KJ, Dawes P, Stewart AJ, Moore DR, Barry JG, et al. Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group 'white paper.' *Int J Audiol*. 2014;53(7):433-45. DOI: 10.3109/14992027.2014.890296
3. Luts H, Eneman K, Wouters J, Schulte M, Vormann M, Buechler M, et al. Multicenter evaluation of signal enhancement algorithms for hearing aids. *J Acoust Soc Am*. 2010;127(3):1491-1505. DOI: 10.1121/1.3299168
4. Ohlenforst B, Zekveld AA, Lunner T, Wendt D, Naylor G, Wang Y, et al. Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation. *Hear Res*. 2017;351:68-79. DOI: 10.1016/j.heares.2017.05.012
5. Anderson Gosselin P, Gagné JP. Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. *Int J Audiol*. 2011;50(11):786-792. DOI: 10.3109/14992027.2011.599870
6. Anderson Gosselin P, Gagné JP. Older Adults Expend More Listening Effort Than Young Adults Recognizing Speech in Noise. *J Speech Lang Hear Res*. 2011;54(3):944. DOI: 10.1044/1092-4388(2010/10-0069)
7. Degeest S, Keppler H, Corthals P. The Effect of Age on Listening Effort. *J Speech Lang Hear Res*. 2015;58(5):1592-1600. DOI: 10.1044/2015_JSLHR-H-14-0288
8. Sarampalis A, Kalluri S, Edwards B, Hafter E. Objective Measures of Listening Effort: Effects of Background Noise and Noise Reduction. *J Speech Lang Hear Res*. 2009;52(5):1230. DOI: 10.1044/1092-4388(2009/08-0111)
9. Picou EM, Ricketts TA, Hornsby BWY. How Hearing Aids, Background Noise, and Visual Cues Influence Objective Listening Effort. *Ear Hear*. 2013;34(5):e52-e64. DOI: 10.1097/AUD.0b013e31827f0431
10. Volz L. Der Unterricht im Absehen der Sprache. *Arch Für Ohren-Nasen- Kehlkopfheilkd*. 1954;165(2-6):362-372. DOI: 10.1007/BF02134821
11. Leonhardt A. Einführung in die Hörgeschädigtenpädagogik: mit 88 Übungsaufgaben und zahlreichen Tabellen. 3., überarbeitete und erweiterte Auflage. München Basel: Reinhardt; 2010.
12. Bernstein JGW, Grant KW. Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*. 2009;125(5):3358-3372. DOI: 10.1121/1.3110132
13. Nirme J, Haake M, Lyberg Åhlander V, Brännström J, Sahlén B. A virtual speaker in noisy classroom conditions: supporting or disrupting children's listening comprehension? *Logoped Phoniater Vocol*. 2018;44(2):79-86. DOI: 10.1080/14015439.2018.1455894
14. Sommers MS, Phelps D. Listening Effort in Younger and Older Adults: A Comparison of Auditory-Only and Auditory-Visual Presentations. *Ear Hear*. 2016;37:62S-68S. DOI: 10.1097/AUD.0000000000000322
15. Krueger M, Schulte M, Brand T, Holube I. Development of an adaptive scaling method for subjective listening effort. *J Acoust Soc Am*. 2017;141(6):4680-4693. DOI: 10.1121/1.4986938
16. Wagener K, Kühnel V, Kollmeier B. Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests. *Z Audiol*. 1999;38(1):4-15.
17. Festen JM, Plomp R. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J Acoust Soc Am*. 1990;88(4):1725-1736. DOI: 10.1121/1.400247
18. Brungart DS. Informational and energetic masking effects in the perception of two simultaneous talkers. *J Acoust Soc Am*. 2001;109(3):1101-1109. DOI: 10.1121/1.1345696
19. Wagener KC, Hochmuth S, Ahrlich M, Kollmeier B. Der weibliche Oldenburger Satztest. In: Deutsche Gesellschaft für Audiologie, ed. Abstracts der 17. Jahrestagung der Deutschen Gesellschaft für Audiologie; 2014. Abrufbar unter/Available from: https://www.dga-ev.com/fileadmin/daten/downloads/bisherige_Jahrestagung/dga2014_programm_final.pdf
20. Llorach G, Kirschner F, Grimm G, Zokoll MA, Wagener KC, Hohmann V. Development and Evaluation of Video Recordings for the OLSA Matrix Sentence Test. *ArXiv E-Prints*. 2019; arXiv:1912.04700
21. Holube I, Fredelake S, Vlaming M, Kollmeier B. Development and analysis of an International Speech Test Signal (ISTS). *Int J Audiol*. 2010;49(12):891-903. DOI: 10.3109/14992027.2010.506889
22. Holube I. Speech intelligibility in fluctuating maskers. *Proceedings of the International Symposium on Auditory and Audiological Research*. 2011;(3):57-64.
23. Oetting D, Brand T, Ewert SD. Optimized loudness-function estimation for categorical loudness scaling data. *Hear Res*. 2014;316:16-27. DOI: 10.1016/j.heares.2014.07.003
24. Koelwijn T, Zekveld AA, Festen JM, Kramer SE. Pupil Dilation Uncovers Extra Listening Effort in the Presence of a Single-Talker Masker. *Ear Hear*. 2012;33(2):291-300. DOI: 10.1097/AUD.0b013e3182310019
25. Rennies J, Schepker H, Holube I, Kollmeier B. Listening effort and speech intelligibility in listening situations affected by noise and reverberation. *J Acoust Soc Am*. 2014;136(5):2642-2653. DOI: 10.1121/1.4897398
26. Schepker H, Haeder K, Rennies J, Holube I. Perceived listening effort and speech intelligibility in reverberation and noise for hearing-impaired listeners. *Int J Audiol*. 2016 Dec;55(12):738-747. DOI: 10.1080/14992027.2016.1219774
27. Schlueter A, Lemke U, Kollmeier B, Holube I. Normal and Time-Compressed Speech: How Does Learning Affect Speech Recognition Thresholds in Noise?. *Trends Hear*. 2016;20:1-13. DOI: 10.1177/2331216516669889
28. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;15(2):155-163. DOI: 10.1016/j.jcm.2016.02.012
29. Mackersie CL, MacPhee IX, Heldt EW. Effects of Hearing Loss on Heart Rate Variability and Skin Conductance Measured During Sentence Recognition in Noise. *Ear Hear*. 2015;36(1):145-154. DOI: 10.1097/AUD.0000000000000091
30. Auer ET, Bernstein LE. Enhanced Visual Speech Perception in Individuals With Early-Onset Hearing Impairment. *J Speech Lang Hear Res*. 2007 Oct;50(5):1157-1165. DOI: 10.1044/1092-4388(2007/080)

Corresponding author:

Saskia Ibelings
 Hörzentrum Oldenburg GmbH, Marie-Curie-Straße 2,
 26129 Oldenburg, Germany, Phone: +49 441 2172-100
 saskia.ibelings@uni-oldenburg.de

Please cite as

Ibelings S, Schulte M, Krüger M, Holube I. Audiovisuelle Realisierung des subjektiven Höranstrengungsmessverfahrens ACALES. *GMS Z Audiol (Audiol Acoust)*. 2020;2:Doc04.
DOI: 10.3205/zaud000008, URN: urn:nbn:de:0183-zaud0000083

This article is freely available from

<https://www.egms.de/en/journals/zaud/2020-2/zaud000008.shtml>

Published: 2020-04-24

Copyright

©2020 Ibelings et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.