

# Individualisierung von Hörunterstützung während eines Telefonats – Experimentelle Untersuchung der kognitiven Belastung und Nutzerakzeptanz

Simon Ciba, Hannah Baumgartner, Tobias Bruns und Jan RENNIES

Fraunhofer IDMT / Hör-, Sprach- und Audiotechnologie, 26129 Oldenburg,

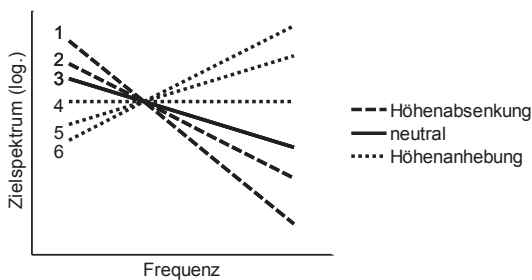
E-Mail: [simon.ciba@idmt.fraunhofer.de](mailto:simon.ciba@idmt.fraunhofer.de)

## Einleitung

Studien haben gezeigt, dass viele Menschen mit Höreinschränkungen grundsätzlich von individualisierter Hörunterstützung in Telefonanwendungen profitieren (z.B. [1,2]). Bisher wurde jedoch noch nicht untersucht, ob eine vom Nutzer selbst durchgeführte Klanganpassung an seine eigenen Bedürfnisse auch während eines Telefongesprächs eingesetzt werden kann oder ob dies zu Problemen führt dem laufenden Gespräch zu folgen. In dieser Studie wurden die bei der Selbstanpassung auftretenden kognitiven Belastungen in einer realitätsnah nachgebildeten Gesprächssituation per Festnetztelefon experimentell erfasst. Außerdem wurde die Zufriedenheit mit den eingestellten Klangeigenschaften mithilfe eines Fragebogens erhoben.

## Preset-basierte Selbstanpassung

Für die Klangindividualisierung wird auf einen Preset-basierten Ansatz [1] zurückgegriffen, welcher bereits in Bezug auf grundlegende Performanzeigenschaften und Gebrauchstauglichkeit untersucht wurde [3,4]. Hinsichtlich technischer Parameter wurde eine für Telekommunikationsanwendungen optimierte Konfiguration gewählt, deren Wirksamkeit zur Verbesserung der Sprachverständlichkeit bereits in [1] gezeigt wurde.

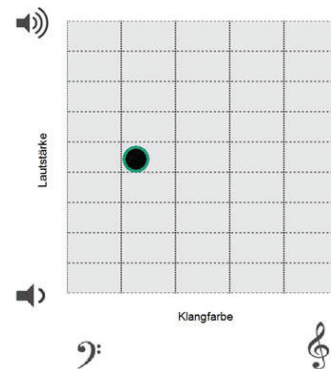


**Abbildung 1:** Schematische Darstellung der Zielspektren des für Telefon optimierten Hörunterstützungsalgorithmus für die sechs verwendeten Preset-Einstellungen.

Der Hörunterstützungsalgorithmus besteht aus einem mehrkanaligen Regelverstärker zur frequenzabhängigen Verarbeitung der klassischen Telefonbandbreite von 300 bis 3400 Hz. In jedem der Frequenzbänder wird das Sprachsignal pegelabhängig verstärkt oder abgeschwächt, so dass das Ausgangssignal in den Bereich vorgegebener Zielwerte verschoben wird. Auf diese Weise werden Schwankungen der Signaleigenschaften ausgeglichen und die Stabilität von Klangbild und Sprachsignalpegel erhöht.

Die frequenzabhängigen Zielpegelwerte wurden nach dem Prinzip einer Klangwaage entwickelt, um – je nach Präferenz des Nutzers – hochfrequente Signalanteile

anzuheben und tieffrequente abzuschwächen oder umgekehrt, und werden vom Nutzer durch Auswahl einer von sechs möglichen Voreinstellungen („Presets“) festgelegt (s. Abbildung 1). Das Zielspektrum von Preset Nr. 3 entspricht dabei einer neutralen Verarbeitung und ist in seiner Form an das ISTS Sprachspektrum im Telefon angelehnt [5].



**Abbildung 2:** Zweidimensionale Webschnittstelle zur Klangindividualisierung durch Positionierung des Cursors: entlang der horizontalen Achse („Klangfarbe“) erfolgt die Variation des Presets, entlang der vertikalen Achse („Lautstärke“) der Verstärkung am Ausgang der Verarbeitung.

Neben der Wahl der Frequenzeinstellung kann der Nutzer die Kennlinien in vertikaler Richtung in Abb. 1 verschieben, was einer linearen Verstärkung bzw. Abschwächung entspricht. Beide Größen werden über eine zweidimensionale Web-Schnittstelle (s. Abbildung 2) durch Positionieren eines Cursors innerhalb der als „Klangfarbe“ und „Lautstärke“ bezeichneten Koordinatenachsen eingestellt.

## Technologieintegration

Kernstück des experimentellen Aufbaus (s. Abbildung 3) ist die Software-Telekommunikationsanlage ASTERISK [6], die eine externe Audiosignalverarbeitung von Telefongesprächen über Ein- und Ausgabe-Ports zum JACK Audio Connection Kit [7] ermöglicht – einem Audio-Server, der es unterschiedlichen Programmen (Clients) erlaubt, über eigene Audio-Ports untereinander Audiosignale auszutauschen. Die Signalverarbeitung wurde in eine C-Bibliothek umgesetzt und als Client in JACK (Hearing Support Algorithm, HSA) integriert. Für die Nutzerinteraktion wurde die oben beschriebene Schnittstelle zur Anpassung der Algorithmusparameter als Web-Interface implementiert. Die eingestellten Parameterwerte werden dem HSA-Client mittels Netzwerkverbindung übermittelt. Bei einem Anruf über einen SIP-Provider verbindet sich die ASTERISK-Telekommunikationsanlage mit dem HSA-

Client. Nach der Verarbeitung wird das Signal über ASTERISK und SIP-Provider an den Hörer eines herkömmlichen Festnetztelefons am fernen Ende weitergeleitet. Auf diese Weise werden alle Strecken realer Telefonverbindungen abgedeckt.

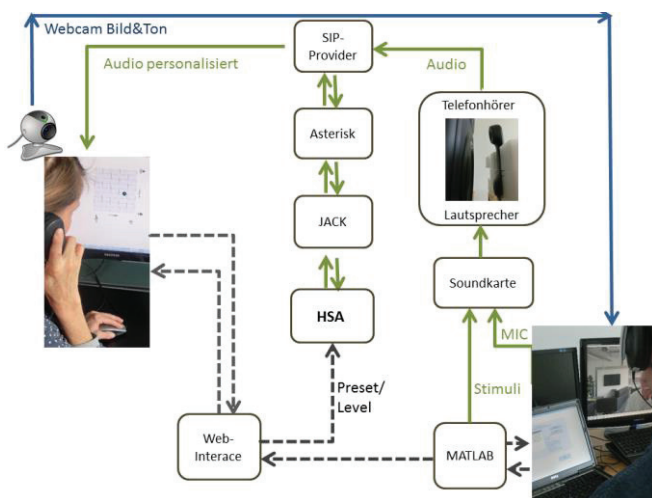


Abbildung 3: Experimenteller Aufbau und Technologieintegration (schematisch).

## Experiment

Neun männliche und acht weibliche Versuchspersonen (VPen) im Alter von 63 bis 81 Jahren mit alterstypischen Hochtonhörverlusten (6 davon Hörgeräteträger) wurden in eine einer Heimumgebung nachempfundene Testumgebung geladen. An einem PC-Arbeitsplatz wurden den VPen über den Telefonhörer Sprachausschnitte vorgespielt, für welche sie den Klang über zuvor beschriebene Webschnittstelle einstellen sollten. Die VPen sollten während des gesamten Experiments jenes Ohr benutzen, welches sie auch im Alltag überwiegend zum Telefonieren einsetzen. Auf diesem „Telefonier-Ohr“ betrug der über die Frequenzen 500 Hz, 1 kHz, 2 kHz und 4 kHz gemittelte Hörverlust (Pure-Tone-Average) der VPen maximal 54 dB HL und durchschnittlich 35 dB HL mit einer Standardabweichung von 9 dB. Der zu Beginn jeder Sitzung erhobene Technikbereitschaftswert nach [8] ergab für die Gesamtheit der VPen im arithmetischen Mittel einen unauffälligen Wert von 3.4 (Median von 3.5).

## Messaufbau und Stimuli

Der Versuchsaufbau (s. Abbildung 3) erforderte die Nutzung zweier Räume. Der Versuchsleiter saß in einem Regieraum, von wo aus er die VP über Webcam sehen und hören konnte. Um die Übertragungseigenschaften einer realen Telefonverbindung nachzubilden, erfolgte die Einspeisung der Testsignale am fernen Ende durch akustische Einkopplung von einem Koaxiallautsprecher in die Sprechmuschel des Telefonhörers, welche in einem Abstand von 6 cm zur Lautsprechermembran fixiert wurde. Die Sprache des Versuchsleiters wurde mit einem Mikrophon abgenommen und über denselben Lautsprecher wiedergegeben, so dass die Möglichkeit zur beidseitigen Kommunikation bestand. Zur Steuerung des Versuchsablaufs zur Laufzeit wurde ein weiteres CGI

implementiert, worüber eine MATLAB-Funktion Befehle sowohl an die Web-Schnittstelle als auch das HSA-PLUGIN senden konnte.

Die Web-Schnittstelle diente der VP zur Klangeinstellung und konnte entweder mittels Monitor und Maus („2DMouse“) oder über ein an den PC angeschlossenes Touch-Screen Gerät („2DTouch“) bedient werden. Beide Eingabemöglichkeiten wurden auf ihre kognitive Last hin getrennt voneinander untersucht.

Als Testsignale wurden den VPen von einem männlichen Sprecher gesprochene Alltagssätze aus dem Inventar des Göttinger Satztests (GÖSA) [9] vorgespielt. Entsprechend den Randbedingungen realer Anwendungen wurde auf eine exakte Kalibrierung des hörunterstützung-internen Pegelmessers verzichtet. Der zugrunde gelegte Übertragungsfaktor von digitalem (dB FS) zu akustischem Pegel (dB SPL) wurde basierend auf Standardübertragungsfaktoren von Telefonsystemen geschätzt. Durch Kunstkopfmessungen wurde überprüft, dass für sämtliche mögliche Parametereinstellungen gesundheitsgefährdende Schalldruckpegel (>110 dB SPL) während des Experiments nicht auftraten.

## Operationalisierung der kognitiven Last

Die VPen wurden aufgefordert, die vorgespielten Sätze zu wiederholen; aus der Anzahl der korrekt wiedergegebenen Wörter im Verhältnis zur Anzahl an Wörtern insgesamt wurde eine „Korrekturrate“ als Maß für die kognitive Leistung errechnet.

Als ein Versuch, den „Flaschenhals“ der Informationsverarbeitung zu verengen und somit in einem mittleren Arbeitsbereich auf der zugehörigen psychometrischen Funktion zu operieren, wurde die Schwierigkeit der Aufgabe erhöht, indem zum einen immer eine Gruppe von mehreren Sätzen auf einmal vorgespielt wurde. Zum anderen wurde einigen der Hörbeispiele ein „Cafeteria-Noise“ als Hintergrundkulisse am fernen Ende additiv beigemischt.

Das Experiment gliederte sich in zwei Sitzungen: einmal die Klanganpassung mit Hilfe der Computermaus und einmal mit Hilfe des Touch-Screens. Die Reihenfolge der Sitzungen innerhalb der Stichprobe wurde ausbalanciert. In beiden Sitzungen wurden die folgenden fünf Versuchsabschnitte durchlaufen:

1. Eingewöhnungsphase: Die VPen machten sich eingangs mit der Bedienung der Web-Schnittstelle vertraut und probierten verschiedene Einstellmöglichkeiten aus. Hierfür wurde ein längerer Informationstext eingespielt. Die VPen konnten diese Aufwärmphase selbständig beenden, sobald sie sich mit der Schnittstelle ausreichend vertraut gemacht hatten.
2. Einstellung des Test-SNR: Um eine evtl. Verschlechterung der kognitiven Leistung möglichst von Effekten peripherer Hörverluste trennen zu können, sollte der Signal-Rausch-Abstand in einem weiteren Schritt so eingestellt werden, dass für die Darbietung im Störgeräusch das subjektiv empfundene Sprachverstehen „ohne

Höranstrengung“ war. Die Signalprobe wurde bei einem anfänglichen SNR von 0dB dargeboten. Entsprechend der Rückmeldung der VP wurde der SNR in 3dB-Schritten solange angehoben, bis ein Sprachverstehen ohne Höranstrengung gegeben war.

3. Referenz-Fit: Den VPen wurden in randomisierter Abfolge 6 Items, bestehend aus jeweils 9 GÖSA-Sätzen vorgespielt. Davon enthielten 3 Items das zuvor beschriebene Störgeräusch. Jeder Satz wurde nur einmal im gesamten Experiment verwendet. Die VPen wurden aufgefordert, während dieser Präsentationen eine für sie im Hinblick auf empfundene Klangqualität und Sprachverständlichkeit insgesamt optimale Einstellung der Hörunterstützung vorzunehmen. Ein Einprägen und Wiederholen der gehörten Sätze war in dieser Kondition nicht erforderlich, die VPen konnten sich ganz auf den Klang konzentrieren. Als Startwert für die Anpassung wurde vom System die leiseste und höhenärmste Parametereinstellung vorgegeben, wodurch eine Bedienung der Schnittstelle durch den VPen in jedem Fall notwendig gemacht wurde. Die Wiedergabe der Sätze lief in Schleife und konnte von den VPen selbstständig beendet werden, so dass ausreichend Zeit für den Einstellungsprozess zur Verfügung stand. Die auf diese Weise gewonnenen Hörunterstützungsparameter wurden als Referenzwerte für eine (subjektiv) als optimal befundene Klangindividualisierung ohne zusätzliche kognitive Last herangezogen.

4. Referenz-Test: Um eine Basisrate für die reine kognitive Leistungsfähigkeit der VP ohne Bedienung einer Schnittstelle zu ermitteln - ein typisches Telefonat kombiniert die Aufgaben Rezeption, Erinnern und Wiedergabe von Sprache – wurden die VPen aufgefordert, sich das Gehörte zu merken und jeweils unmittelbar nach der Darbietung eines Items zu wiederholen. Erneut wurden 3 Items mit und 3 Items ohne Störgeräusch in randomisierter Abfolge dargeboten. Jedes Item enthielt dabei 3 Sätze aus dem GÖSA, wobei die Wiedergabe automatisch nach einmaligem Durchlauf beendet wurde. Aus den korrekt erinnerten Wortanteilen wurden die eingangs beschriebenen Korrektraten berechnet.

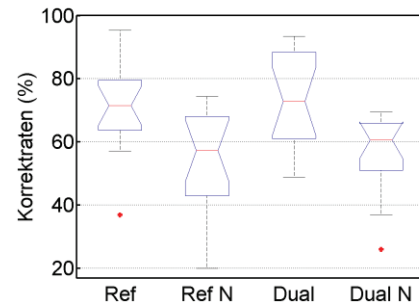
5. Dual-Task-Test: Auch im letzten Abschnitt wurden jeweils 3 Items von je 3 Sätzen sowohl mit als auch ohne Störgeräusch in randomisierter Abfolge vorgespielt. Wie bereits im Referenz-Fitting wurden die VPen aufgefordert, die dargebotenen Sätze in ihrem Klang zu optimieren. Zusätzlich sollte unmittelbar im Anschluss an die einmalige Darbietung der Sätze deren Inhalt wiederholt werden (wie in Referenz-Test).

In der Auswertung wurden die Abweichungen der in dieser „Dual-Task“-Situation eingestellten Hörunterstützungsparameter und der gemessenen Korrektraten zu den jeweiligen Referenzkonditionen analysiert.

Im Anschluss an jede Sitzung wurde der System-Usability-Scale (SUS) Wert nach [10] für die jeweilige Eingabevariante erhoben. Mit einem weiteren, eigens konzipierten Fragebogen wurden Aspekte der Nutzerzufriedenheit untersucht.

## Ergebnisse

Abbildung 4 zeigt die Häufigkeitsverteilungen der für die 2D-Mouse-Bedienvariante unter der Referenz- und Dual-Task-Bedingung ermittelten Korrektraten als Boxplots, jeweils bei vorhandener und fehlender Hintergrundkulisse am fernen Ende. Jeder Messpunkt wurde aus Zählung der Wörter über alle drei Items ermittelt. Durch die Störung werden die mittleren Korrektraten erheblich verschlechtert (Absinken des Medians von mehr als 10%). Das Verteilungsprofil für die 2DTouch-Bedienung (ohne Abbildung) weist nur geringfügige, nicht signifikante Unterschiede zum dargestellten auf.



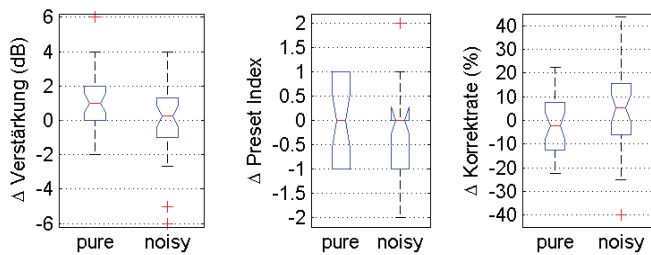
**Abbildung 4:** Korrektraten für Referenz- („Ref“) und Dual-Task-Test („Dual“) – jeweils mit („N“) und ohne Hintergrundkulisse; Darstellung für 2D-Mouse-Bedienvariante.

In den Verteilungen der gewählten Presets und Verstärkungsstufen (ohne Abbildung) weichen die Mediane beider Bedienvarianten ebenfalls nur leicht und nicht signifikant voneinander ab. Aus einer zweidimensionalen Häufigkeitsanalyse aller gewählten Kombinationen aus Preset und Verstärkung geht außerdem hervor, dass der zu Verfügung gestellte Parameterraum bis auf Randbereiche (Preset 6, Verstärkungsstufe 1) umfänglich genutzt wurde.

Um die Unterschiede der Messgrößen zwischen Referenz-Bedingung (Referenz-Fit bzw. Referenz-Test) und Dual-Task-Bedingung, genauer untersuchen zu können und hierfür die intersubjektvariabilität der Daten zu eliminieren, wurden für jede VP die Differenzen der Messgrößen zwischen beiden Bedingungen ermittelt. Für die eingestellten Parameter wurden hierbei Mittelwertschätzer (arithmetisches Mittel für Verstärkung in dB, Median für Preset-Index) verwendet, um die Messwerte aller 3 Items zusammen zu fassen. Abbildung 5 zeigt die Boxplots der Differenzen sowohl in Ruhe als auch vor Hintergrundrauschen.

In Ruhe („pure“) deuten ein Median von ca. +1dB und eine die Nulllinie nicht überlappende Einkerbung des Boxplots auf eine leichte, jedoch bereits signifikante ( $\alpha = 5\%$ ) Erhöhung der Abhörlautstärke in der Dual-Task Situation hin. Hingegen erfolgt die Wahl der Presets bei einer mittleren Differenz von 0 und symmetrischen Boxplots konditionsunabhängig. Die Korrektrate sinkt im Median um 2.3% leicht ab. Bei vorhandener Hintergrundkulisse („noisy“) kann im Mittel nur eine sehr geringe Erhöhung der Abhörlautstärke von <0,5 dB gegenüber der Single-Task Situation beobachtet werden. Für die Presets beträgt der

Median der Änderung auch hier Null. Allerdings deutet ein negativer Schwerpunkt der Verteilung tendenziell auf die Wahl von höhenärmeren Presets hin. Für die Korrektraten kann nun eine signifikante Zunahme, bei einem Median von 5.2 %, beobachtet werden.



**Abbildung 5:** Differenzen der drei experimentellen Messgrößen zwischen Dual-Task- und Referenz-Bedingung (positive Differenzen bedeuten höhere Werte im Dual-Task); Boxplots sind jeweils gruppiert nach dem Vorhandensein von Hintergrundgeräusch („pure“/„noisy“) und basieren auf den Daten beider Sitzungen (2DMouse/2DTouch).

Der aus den Fragebögen resultierende SUS-Wert beträgt für die 2DTouch-Bedienung im arithmetischen Mittel 76,5 und im Median 77.5 und liegt somit nur geringfügig höher als für das 2DMouse mit 71.3 bzw. 75. Die Bewertung der Zufriedenheit mit dem Ergebnis der Klangindividualisierung (unabhängig von der Bedienvariante) erfolgte auf einer fünfstufigen Ratingskala mit den Kategorien von „deutlich besser“, über „etwas besser“, „unverändert“, „schlechter“ bis „deutlich schlechter“. Die Häufigkeitsverteilung der gewählten Kategorien (in Reihenfolge der Aufzählung) beträgt für die Kategorie „Klang“ 7-7-3-0-0, für „Lautstärke“ 7-5-5-0-0 und für „Sprachverständlichkeit“ 8-6-3-0-0. Unter den Höreräteträgern ergaben sich für die Kategorie „Klang“ die Häufigkeiten 3-2-1-0-0, für „Lautstärke“ 3-1-2-0-0 und für „Sprachverständlichkeit“ 4-1-1-0-0.

## Diskussion und Fazit

Die Ergebnisse dieser Studie zeigen, dass auch in einer realitätsnahen Kommunikationssituation bei Verwendung einer geeigneten Schnittstelle eine individuell einzustellende Hörunterstützung für ältere schwerhörnde Nutzer möglich ist. Es zeigten sich dabei zwischen den Bedienvarianten 2DMouse und 2DTouch hinsichtlich der untersuchten Messgrößen keine wesentlichen Unterschiede, was darauf hindeutet, dass sowohl PC- als auch Tablet / Smartphone-Lösungen verwertbar sind.

Die experimentellen Befunde deuten an, dass bei ruhigem Hintergrund am fernen Ende das zusätzliche Fitting während des Zuhörens gegenüber der seriellen Abarbeitung der beiden Aufgaben die Performanz der VP im „Merktest“ und somit die Fähigkeit einem Telefongespräch inhaltlich zu folgen nur sehr geringfügig (und hier nicht signifikant) verschlechtert. Bei widrigen Übertragungsbedingungen, wie etwa bei der hier hinzugefügten Hintergrundkulisse, kann durch die Möglichkeit der Klangveränderung „in situ“, d.h. während der Leistungsaufgabe, gegenüber der Single-Task-Bedingung die Performanz insgesamt sogar verbessert werden.

Neben den experimentellen Daten ergaben die subjektiven Bewertungen der VPen eine hohe Akzeptanz der selbstgesteuerten Hörunterstützung am Telefon. Dies war gleichermaßen für Höreräteträger und unversorgte Schwerhörnde der Fall, so dass davon auszugehen ist, dass die erforschten Ansätze Potenzial zur Verwertung in beiden Nutzergruppen haben.

## Danksagung

Diese Arbeit wurde durch das Bundesministerium für Bildung und Forschung (Förderkennzeichen 16SV6182) gefördert. Wir danken außerdem unseren Kollegen Sylvain Favrot, Yves Wagner, Richard Paluch, Henning Küwen und Dirk Oetting für ihre Mitarbeit und Unterstützung.

## Literatur

- [1] Oetting, D., Rannies, J. & Appell, J. E. (2014). Hearing supportive algorithm with self-adjustment user interface integrated in the telephone network. In: World Telecommunications Congress 2014, 1-5.
- [2] Roup, C. M., Poling, G. L., Harhager, K., Krishnamurthy, A., & Feth, L. L. (2011). Evaluation of a Telephone Speech-Enhancement Algorithm Among Older Adults With Hearing Loss. *J. Speech Lang. Hear. R.* 54(5), 1477-1483.
- [3] Ciba, S., Volgenandt, A., Bruns, T., Asendorf, R., Oetting, D., Rannies, J. (2014): Evaluation of interfaces for the self-fitting of personalized communication systems by hearing-impaired users. *Fortschritte der Akustik, DAGA 2014*, 68-69.
- [4] Baumgartner, H., Paluch, R., Fuhrmann, K., Rannies, J., Meis, M., Appell, J. E. (2014): Usability Evaluation of Self-Fitting Interfaces for Personalized Sound Systems. *Akustik & Audiologie – Hören für alle, DGA 2014*.
- [5] Byrne, D., Dillon, H., Tran, K. et al. (1994). An international comparison of long-term average speech spectra. *J. Acoust. Soc. Am.* 96(4), 2108-2120.
- [6] <http://www.asterisk.org/>
- [7] <http://jackaudio.org/>
- [8] Neyer, F. J., Felber, J., & Gebhardt, C. (2012). Entwicklung und Validierung einer Kurzskaala zur Erfassung von Technikbereitschaft. *Diagnostica*, 58(2), 87-99.
- [9] Kollmeier, B., Wesselkamp, M., (1997): Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment. *J. Acoust. Soc. Am.* 102, 2412-2421.
- [10] Brooke, J. (1986): SUS: A “quick and dirty” usability scale. In: Jordan, P. W., Thomas, B., Weerdmeester, B., McClelland (Hrsg.) *Usability Evaluation in Industry*, 189—194, Taylor & Francis, London, UK.