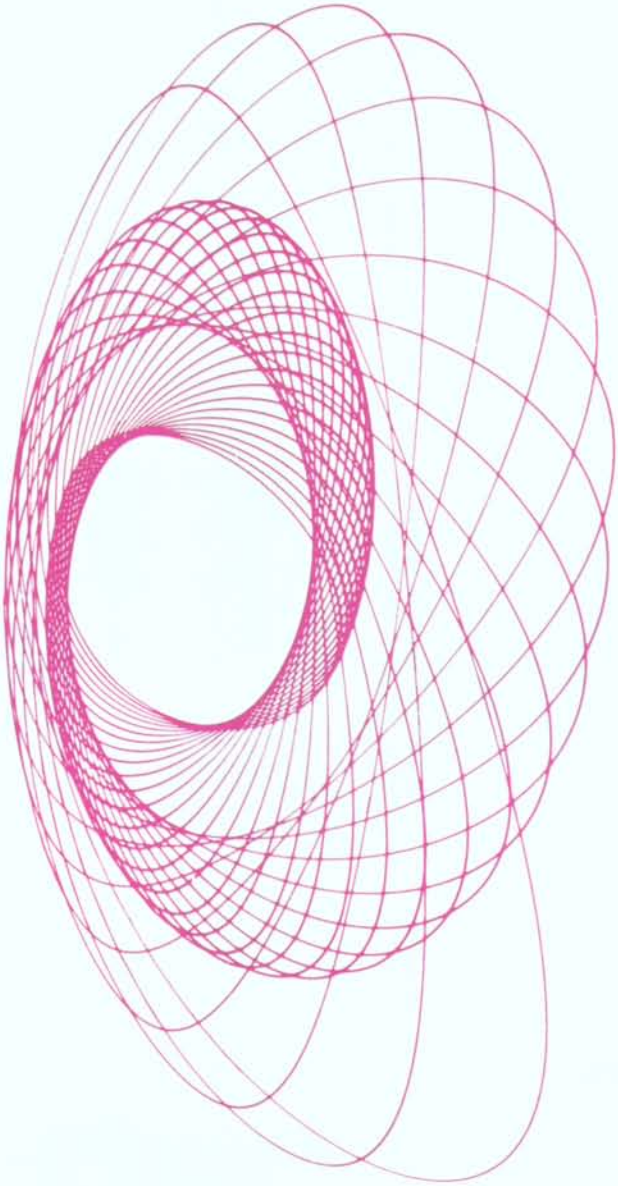
D 21976

Zeitschrift für Audiologie

Audiological Acoustics

ISSN 1435-4691



Editorial

»Zeitschrift für Audiologie« am Scheideweg
J. Kießling (Gießen)

Originalarbeiten

Noise reduction schemes for digital hearing aids –
Part I: Listening effort and speech intelligibility
M. Matzinzik, B. Kollmeier (Oldenburg)

Evaluation of different 3-channel dynamic compression schemes
in a field test with wearable DSP prototype hearing aids
J.-E. Appell, V. Hohmann, B. Gabriel, B. Kollmeier (Oldenburg)

Mitteilungen der Deutschen Gesellschaft für Audiologie

Jahrgang 41 Nr. 1 2002

Zeitschrift für Audiologie

Audiological Acoustics

Mitteilungen der
Deutschen Gesellschaft für Audiologie

Impressum

Verlag und Redaktion:
Median-Verlag
von Killisch-Horn GmbH
Hauptstraße 64
69117 Heidelberg
Postfach 10 39 64
69029 Heidelberg
Telefon (0 62 21) 9 05 09-0
Fax (0 62 21) 9 05 09-20
E-mail:
median-verlag@t-online.de

Redaktion:
Christina Osterwald
Anzeigen:
Karin Ball
Layout:
Median-Verlag GmbH
Druck:
Dietz Druck, Heidelberg

ISSN 1435-4691

Erscheinungsweise:
viermal jährlich
jeweils Mitte des Quartals
Bezugspreis jährlich:
EURO 35,00
SFr 62,00
zuzüglich Versandkosten:
Inland EURO 6,00
Ausland EURO 13,00
Luftpost EURO 23,00

Zur Zeit hat Anzeigenliste Nr. 12
vom 1. Januar 2002 Gültigkeit.

Publisher:
Median-Verlag
von Killisch-Horn GmbH
Hauptstraße 64
D-69117 Heidelberg
P. O. Box 10 39 64
D-69029 Heidelberg
Phone (+49 62 21) 9 05 09-0
Fax (+49 62 21) 9 05 09-20
E-mail:
median-verlag@t-online.de

Editorial department:
Christina Osterwald
Advertising:
Karin Ball
Layout:
Median-Verlag GmbH
Printed
by Dietz Druck, D-Heidelberg

Published quarterly
4 issues per annum
Annual subscription rates:
EURO 35,00
SFr 62,00
plus surface mail:
EURO 13,00
or air mail:
EURO 23,00

Supplied directly by
Median-Verlag GmbH
P. O. Box 10 39 64
D-69029 Heidelberg
Current advertisement rates
No. 12, January 1st, 2002.

Z Audiol
41 (1) 1–46 (2002)

Zeitschrift für Audiologie

Audiological Acoustics

Editorial

1 **»Zeitschrift für Audiologie« am Scheideweg**
Jürgen Kießling (Gießen)

Originalarbeiten

4 **Nose reduction schemes for digital hearing aids – Part I: Listening effort and speech intelligibility**
Störgeräuschunterdrückungsalgorithmen für digitale Hörgeräte – Teil I: Zuhöranstrengung und Sprachverständlichkeit
Mark Marzinzik, Birger Kollmeier (Oldenburg)

22 **Evaluation of different 3-channel dynamic compression schemes in a field test with wearable DSP prototype hearing aids**
Feldtest verschiedener Dreikanal-Dynamikkompressionsalgorithmen mit einem tragbaren digitalen Hörgeräteprototypen
Jens-E. Appell, Volker Hohmann, Birgitta Gabriel, Birger Kollmeier (Oldenburg)

A2 ***Autorenhinweise***

A9 ***Veranstaltungen***

Autorenhinweise

Allgemeines

Die »Zeitschrift für Audiologie« veröffentlicht Arbeiten auf dem Gebiet der Audiologie und Neurootologie sowie, sofern sie von Interesse für den Leserkreis sind, auf allen Gebieten, die eine unmittelbare oder mittelbare Beziehung zur Audiologie haben wie Phoniatrie und Pädaudiologie, Anatomie und Physiologie, Biochemie und Pharmakologie, Molekularbiologie und Genetik, Arbeitsmedizin und Epidemiologie, Biologie, Psychologie und Pädagogik, Phonetik, Logopädie und Kommunikationswissenschaften, Signalerkennungstheorie und Psychophysiologie, Physik, Akustik (insbesondere Psychoakustik und Elektroakustik), Elektrotechnik und Hörgeräte-Akustik.

Einreichen von Manuskripten

Nur Originalarbeiten in deutscher oder englischer Sprache sind zu senden an:

Univ.-Prof. Dr. rer. nat. Jürgen Kießling
Hals-Nasen-Ohrenklinik
Justus-Liebig-Universität Gießen
Feulgenstraße 10
35385 Gießen
Tel.: +49-641-99-43790
Fax: +49-641-99-43799
E-mail: juergen.kiessling@hno.med.uni-giessen.de

• Die Manuskripte sind vierfach (einschließlich aller Abbildungen) einzureichen. Kopien von Mikrofotographien sowie Schwarzweißkopien von Farbabbildungen können nicht akzeptiert werden. Die Blätter sind einseitig mit doppeltem Zeilenabstand zu beschreiben; die Breite des linken Randes sollte etwa 3 cm betragen.
• Mit dem revidierten Manuskript ist auch eine Diskette mit Text, Tabellen und Legenden (Word für Windows 6.0, 7.0, 97 oder kompatibles Programm) sowie Abbildungen (bevorzugte Formate .crd, .bmp, .jpg, .eps oder .tif) einzureichen.
• Die Autoren erhalten fünf Belegexemplare des Heftes sowie 50 Sonderdrucke frei. Mit der Korrekturfahne werden ein Bestellvordruck und eine Preisliste für Sonderdrucke verschickt. Für Sonderdrucke, die erst nach dem Druck des Heftes bestellt werden, wird ein höherer Preis erhoben.
• Manuskripte, die zur Veröffentlichung eingereicht werden, dürfen nicht gleichzeitig anderen Verlagen und Redaktionen zur Publikation angeboten werden. Die Autoren tragen die Verantwortung dafür, daß die eingereichte Arbeit noch nicht anderweitig mit gleichem oder ähnlichem Inhalt veröffentlicht worden ist und daß sie das Urheberrecht besitzen. Auch obliegt es den Autoren, die Erlaubnis zur Reproduktion von Abbildungen, Tafeln usw. aus anderen Publikationen einzuholen.
• Manuskripte können als Originalarbeiten oder als Kurzmitteilungen eingereicht werden. Für unverlangt eingesandte Übersichtsarbeiten sowie Fort- und Weiterbildungsartikel wird keine Garantie übernommen; es besteht allerdings die Chance, daß sie nach Überprüfung durch zuständige Experten veröffentlicht werden.
• Manuskripte von Originalarbeiten sowie Fort- und Weiterbildungsartikel unterliegen der Begutachtung durch mindestens zwei unabhängige Experten. Kurzmitteilungen (vorläufige Ergebnisse) werden im wesentlichen nur nach formalen Kriterien überprüft.
• Im Falle einer Ablehnung des Manuskripts wird den Autoren die Möglichkeit eingeräumt, ihr Manuskript innerhalb einer Frist von drei Monaten zu revidieren. Die revidierten Manuskripte unterliegen einer erneuten Begutachtung. Die Ablehnung des revidierten Manuskripts ist endgültig.

Aufbau der Arbeiten

Titelseite:

Titel der Arbeit (deutsch und englisch, eventuell mit Untertitel, der als solcher zu kennzeichnen ist);
Autorenname(n) und nicht abgekürzte(r) Vorname(n);
Institut und Ort, Land;
Kurztitel (sog. »lebender Kolumnentitel«);
Key words (deutsch und englisch);

Guidelines for authors

General information

The *Zeitschrift für Audiologie* (Audiology Journal) publishes papers on audiology and neurootology in addition to contributions on areas directly or indirectly associated with audiology if these are of general interest to its readers. Such areas include phoniatry and pedaudiology, anatomy and physiology, biochemistry and pharmacology, molecular biology and genetics, occupational health and epidemiology as well as biology, psychology, education theory, phonetics, speech therapy and communication studies, signal recognition theory, psychophysiology, physics, acoustics (especially psychoacoustics and electroacoustics), electrical engineering and hearing-aid acoustics.

Submitting manuscripts

We only accept original work and this must be either in German or English. Contributions should be sent to the following address:

Prof. Dr. Juergen Kiessling
Ear, Nose and Throat Department
The Justus Liebig University of Giessen
Feulgenstrasse 10
D – 35385 Giessen (Germany)
Phone: +49-641-99-43790
Fax: +49-641-99-43799
E-mail: juergen.kiessling@hno.med.uni-giessen.de

• Manuscripts must be submitted in quadruplicate (including all illustrations). Microphotography and black and white copies of colour illustrations are not accepted. The manuscripts must be typed on single sheets using double line spacing. The left-hand margin should be approx. 3 cm.
• The text, tables and captions of revised manuscripts should also be submitted on 3.5" diskette (Word for Windows 6.0, 7.0, 97 or a compatible program). This also applies to any graphics (preferred formats: .crd, .bmp, .jpg, .eps or .tif).
• Authors will receive free of charge five copies of the issue in which their contribution appears as well as 50 offprints. An order form and the offprint price list are included with the galley proofs. A higher price is charged if offprints are ordered after going to press.
• Manuscripts being submitted for publication may not be offered to other publishers. It is the responsibility of the author to ensure that the paper being submitted has not been published in similar form elsewhere and that they hold the copyright. It is also the author's responsibility to apply to the relevant sources for permission to reproduce illustrations, tables etc. from other publications.
• Authors may only submit original work or preliminary reports. Previews (summaries) and educational articles may only be submitted if specifically requested by the editors.
• Original work and educational articles will be subject to the opinion of at least two independent experts. Previews (preliminary results) are generally only checked for form.
• Should a manuscript be rejected, the author has the option of revising and resubmitting it within three months. Revised manuscripts will also be monitored by experts. The decision to reject a revised manuscript is final.

Manuscript layout

Title page

Title of work in English and German (if there is a subtitle, this must be marked as such)
Name(s) of author(s) and full first name(s)
Name of institute, city and country of location
Short title (column heading)
Key words (in English and German)
Official contact address for correspondence and mailing offprints (with telephone / fax numbers and E-mail address)

2nd page

Abstract in the same language as the manuscript, maximum 1 printed page

Vollständige Postanschrift des für Korrespondenz und Sonderdruckversand zuständigen Autors (einschließlich Telefon- und Telefaxnummern sowie E-mail-Adresse).

2. Seite:

Abstract von maximal einer Druckseite in jener Sprache, in der das Manuskript abgefaßt ist.

3. (und 4.) Seite:

Extended Abstract von maximal zwei Druckseiten (Originalarbeiten) bzw. einer Druckseite (Kurzmitteilungen) in der jeweils anderen Sprache

Folgende Seiten

Originalarbeiten inklusive Tabellen und Abbildungen sollten nicht mehr als zehn Druckseiten umfassen, Kurzmitteilungen nicht mehr als fünf Druckseiten:
Einleitung; Material und Methodik bzw. Krankengeschichten;
Ergebnisse; Diskussion; Zusammenfassung (fakultativ); Literatur;
Tabellen mit Titeln; Abbildungen mit Legenden; eventuell Anhänge.

Kleinschrift

Textpartien, die in kleinerer Schrift gesetzt werden sollen, sind am linken Rand mit einem senkrechten Strich und einem »P« (für »petit«) zu kennzeichnen und ebenfalls mit doppelter Zeilenschaltung zu schreiben.

Auszeichnungen, Hervorhebung durch andere Schrift

Hervorhebungen im Text sollen sparsam angewendet werden. Einzelne Wörter oder Satzteile, die der Autor hervorzuheben wünscht, sollen entsprechend gekennzeichnet werden; sie werden *kursiv* gesetzt. Die halbfette Schrift ist für bestimmte Untertitelabstufungen sowie Tabellen- und Abbildungsnummern vorbehalten; für den laufenden Text steht sie nicht zur Verfügung. Nicht ausgezeichnet werden lateinische Begriffe, wie »in vivo«, »in vitro«, »et al.«, »in utero«, »pars pro toto« usw. Lateinische botanische und zoologische Gattungsnamen sind vom Autor zu kennzeichnen; sie werden kursiv gesetzt (*Echinus esculentus linnaeus*).

Fußnoten

Fußnoten sind auf ein Minimum zu beschränken, weil sie den Lesefluß hemmen. Die meisten Angaben in Fußnoten können ohnehin im Text untergebracht werden, z. B. in Klammern oder als Kleinschriftabsätze.

Tabellen

Resultate, die in Tabellenform präsentiert werden, sollten nicht gleichzeitig im Text wiederholt und außerdem grafisch dargestellt werden. Alle Tabellen werden pro Arbeit mit 1 beginnend durchgehend arabisch numeriert. Jede Tabelle benötigt einen Titel. Im Text wird jede Tabelle chronologisch und an der Stelle erwähnt, wo sie im fertigen Umbruch ungefähr stehen soll.

Tabellen sind mit doppelter Zeilenschaltung auf separate Blätter zu schreiben, am Schluß des Manuskripts beizulegen und wie die übrigen Manuskriptseiten mitzunumerieren. Der Verlag hat für die einheitliche typographische Darstellung der Tabellen Richtlinien entworfen, die verbindlich sind. Fußnoten und Bemerkungen zum Tabelleninhalt werden in die Tabelle einbezogen. Als Fußnotenhinweise dienen normalerweise hochgestellte arabische Ziffern, pro Tabelle immer mit 1 beginnend. Besteht Verwechslungsgefahr mit den in der Tabelle vorkommenden Potenzziffern, werden als Fußnotenhinweise anstelle der Ziffern hochgestellte Kleinbuchstaben verwendet.

Schwarzweiß-Abbildungen

Abbildungsvorlagen sind auf der Rückseite mit einem weichen Bleistift und der gebotenen Vorsicht mit der Abbildungsnummer, dem Namen des Autors und der Angabe, wo auf der Abbildung »oben« ist, zu versehen.

Alle Textangaben in Abbildungen müssen grundsätzlich in englisch erfolgen. Abbildungsvorlagen mit Beschriftungen aus Fremdquellen sind ins Englische zu übersetzen.

Alle Abbildungen werden pro Arbeit mit 1 beginnend durchgehend arabisch numeriert. Im Text wird jede Abbildung chronologisch und an der Stelle erwähnt, wo sie im fertigen Umbruch ungefähr stehen soll.

Für die Wiedergabe von Fotografien sind Hochglanzabzüge erforderlich; Röntgenbilder können als verkleinerte Papierkopien vorgelegt werden.

Für Darstellungen in Kurven und Grafiken empfehlen wir, einfache geometrische Symbole zu verwenden.

Farbige Abbildungen

Als Vorlagen für Farbabbildungen sind entweder Hochglanzabzüge oder Diapositive erforderlich.

Abbildungslegenden

Jede Abbildung benötigt eine Legende. Sämtliche Abbildungslegenden (eben

3rd (and 4th) page

Extended abstract in the other language, maximum length: 2 printed pages for original work, 1 printed page for previews

Following pages

Introduction
Materials and method; case histories, if applicable
Results
Discussion
Conclusion (optional)
Bibliography
Tables with titles
Illustrations with captions
Appendices (if applicable)
(Note: Original work should be no longer than ten printed pages, including all tables and illustrations. Previews should not exceed five printed pages)

Small type

Parts of the text to be set in small type should be indicated in the left margin by a vertical line and the letter »P« (for »petit«). This text should also be double spaced.

Highlighted text, different type styles

Authors should avoid highlighting the text wherever possible. They should use a single underline to emphasize or highlight individual words or parts of sentences. These will then be typeset in *italics*. Secondary bold type may not be used in the body of the text, as it is reserved for certain types of subtitles and for numbering tables and illustrations. Common Latin expressions such as »in vivo«, »in vitro«, »et al.«, »in utero«, »pars pro toto« etc. are not highlighted or italicized. Authors should underline any botanical and zoological terms that appear in Latin. These will also be typeset in italics (*Echinus esculentus linnaeus*).

Footnotes

Footnotes should be kept to a minimum because they interrupt the flow of the text for the reader. The contents of most footnotes can be integrated into the text itself by using brackets or separate paragraphs in small type.

Tables

Results presented in table form should not be repeated in the text or as graphics. All tables in a paper must be numbered consecutively using Arabic numerals. Each table should also be given a title. The approximate position of the tables (in order of appearance) in the finished text must be indicated in the manuscript.

Tables must also be double spaced and a separate page used for each table. These pages are added at the end of the paper and should follow the consecutive numbering used in the manuscript. We have developed a standardized typography for all tables and our guidelines are binding. Footnotes and comments on table content should be incorporated into the table itself. Footnotes are referenced in superscript by Arabic numerals and are numbered separately for each table, starting with 1. If there is a danger of these numerals being confused with the numbers indicating the mathematical power of the values given in the table, lowercase letters of the alphabet should be substituted as the superscript for footnote references.

Black and white illustrations

The number of the relevant illustration and the author's name should be lightly pencilled on the back of each illustration, along with some indication of which way up the illustration is to be printed. This is best done with a soft lead pencil.

All texts in the illustrations must be in English. Where illustrations from other sources are used, any text not already in English must be translated.

All illustrations in any one paper are to be numbered consecutively using Arabic numerals. The approximate position of the illustrations (in order of appearance) in the finished text must be indicated in the manuscript.

Reproductions of photographs can only be made from gloss prints. Photocopies of X-rays can be printed if these have been reduced in size.

Markings on graphs and graphics should be in the form of simple geometric symbols.

Colour illustrations

Colour illustrations can only be printed from gloss prints or from transparencies.

Captions for illustrations

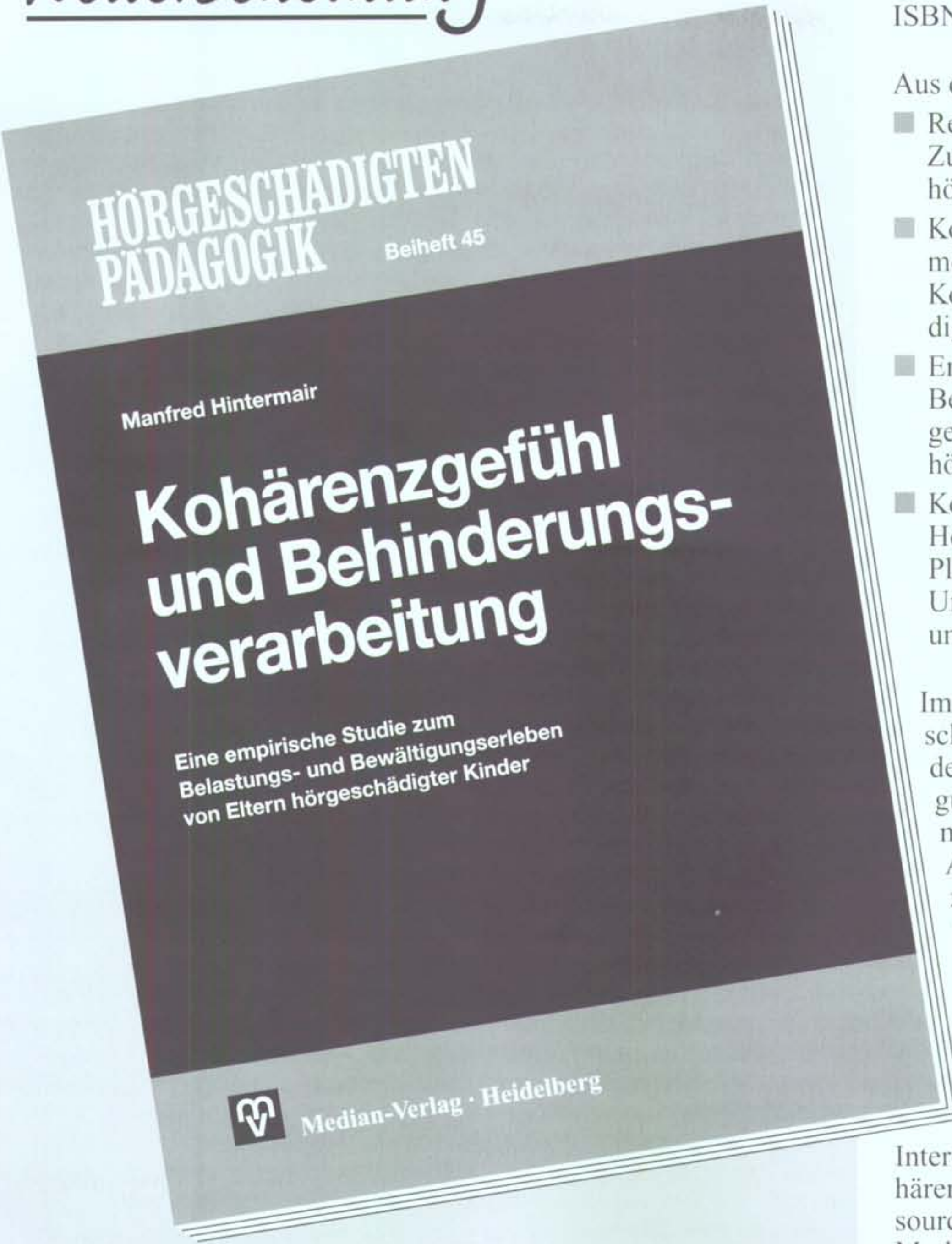
Each illustration – and each table – must be given a caption in both English and German. The captions should be numbered consecutively and listed (using double spacing) on a separate page at the end of the manuscript.

AUTORENHINWEISE
GUIDELINES FOR AUTHORS

Abbildungslegenden
Jede Abbildung – und jede Tabelle – benötigt eine Legende. Sämtliche Abbildungslegenden (ebenfalls mit doppeltem Zeilenabstand) sind zweisprachig (deutsch und englisch) zu verfassen und am Schluß des Manuskripts beizulegen.
Orthographie, Grammatik, sprachliche Korrektur
Maßgebend für die Orthographie und Grammatik sowie die medizinische Terminologie der deutschsprachigen Publikationen sind die verschiedenen Bände des »Großen Duden«, der »Medizin-Duden« und das »Klinische Wörterbuch« von Pschyrembel.
In deutschsprachigen Manuskripten ist die eindeutschende Schreibung lateinischer und griechischer Begriffe die Regel (Kalzium anstatt Calcium; Kortex anstatt Cortex). Für englischsprachige Arbeiten halte man »Webster's New Collegiate Dictionary« bzw. »Oxford English Dictionary«. Die Schreibweise medizinischer Fachbegriffe richtet sich nach dem »Stedman's Medical Dictionary« oder dem »Dorland' Illustrated Medical Dictionary«.
Den Autoren wird empfohlen, Textteile, die nicht in ihrer Muttersprache geschrieben sind, einer sprachlich kompetenten Person mit entsprechender Muttersprache zur sprachlichen Korrektur vorzulegen. Sie werden zusätzlich von fachkundigen native speakers überarbeitet.
Terminologie
Die Autoren sind gehalten, international empfohlene Terminologien zu benutzen (Empfehlungen der ISO-IEC, Nomina Anatomica, WHO List of Approved Names for Drugs). Die Benennung von Einheiten muß der internationalen Norm («Système International d'Unités», SI) entsprechen.
Abkürzungen
Der Gebrauch von Abkürzungen sollte auf ein Mindestmaß reduziert werden; ihre Bedeutung muß bei ihrer ersten Verwendung deutlich erklärt werden.
Audiogramme
Alle Audiogramme müssen entsprechend den ISO-Standards gezeichnet werden.
Korrekturfahne
Die Korrekturfahne wird dem korrespondierenden Autor vorgelegt und sollte umgehend zurückgesandt werden.
Referenzen
Literaturhinweise im Text erfolgen (in Klammern) durch Nennen des Namens des Verfassers (*kursiv*) und des Erscheinungsjahres. Ein Autor: (*Näätänen*, 1994); zwei Autoren: (*Kraus* und *MacGee*, 1994); drei Autoren: (*Picton* et al., 1994). Zur Publikation eingereichte, aber noch nicht akzeptierte Arbeiten sind mit ›unpubliziert‹ zu bezeichnen und nicht in das Literaturverzeichnis aufzunehmen. Für die Zeitschriften sind die Abkürzungen des Index Medicus zu verwenden.
Das Literaturverzeichnis soll ausschließlich im Text zitierte Publikationen enthalten. Namen und Initiale(n) der Autoren (ohne Punkt) sollen nicht durch Kommata getrennt werden; nur verschiedene Autoren sind durch Kommata zu trennen. Alle Autoren müssen aufgeführt werden; ›et al.‹ ist unzureichend. Werden von den gleichen Autoren mehrere Arbeiten aus dem gleichen Jahr zitiert, ist hinter der Jahreszahl der Buchstabe a, b, c usw. anzufügen, z. B. (*Kraus*, 1994a).
Beispiele:
In *Zeitschriften* veröffentlichte Arbeiten: Näätänen R, Picton T (1987) The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. Psychophysiology 24, 375–425
Monographien: Näätänen R (1992) Attention and Brain Function. Lawrence Erlbaum Associates, Hillsdale, NJ.
Buchbeiträge: Kraus N, McGee T (1994) Auditory event-related potentials. In: Katz J (Hrsg.) Handbook of Clinical Audiology. Williams & Wilkins, Baltimore, Hongkong, London, München, Sydney, Tokyo, S 403–423.
Curriculum vitae
Für das Curriculum vitae werden Portraitfotos (schwarzweiß oder farbig) sowie kurze Lebensläufe aller Autoren in deutscher und englischer Sprache mit vollständiger (ggf. auch E-mail) Adresse erbeten.

Spelling, grammar and correct usage
The spelling, grammar and medical terminology used for work submitted in German are based on the following works of reference: *Der Große Duden* (all volumes), *Der Medizin-Duden* and *Das Klinische Wörterbuch* (Pschyrembel). The relevant works of reference for basic English usage are *Webster's New Collegiate Dictionary* or *The Oxford English Dictionary*. English medical terminology is based on *Stedman's Medical Dictionary* or *Dorland's Illustrated Medical Dictionary*.
Latin and Greek terms used in German papers should follow the accepted Germanized form (Kalzium and Kortex, not Calcium and Cortex). For the correct English usage, please consult the works of reference given above.
We strongly recommend that authors submitting texts or parts of texts not written in their mother tongue have these pre-edited by a linguistically competent native speaker of the language in question.
Terminology
Authors are required to use the accepted international terminology (Recommendations of the ISO-IEC, Nomina Anatomica, WHO List of Approved Names for Drugs). Names of units must comply with the international norms (*Système International d'Unités, SI*).
Abbreviations
Abbreviations should be avoided if at all possible. If used, the full meaning should be clearly stated when the abbreviation first appears in the text.
Audiograms
All drawings of audiograms must comply with ISO standards.
Galley proofs
Galley proofs are sent to the contributor's official contact address for correspondence and should be returned without delay.
References
References in the text to secondary literature are placed in brackets and give the name(s) of the author(s) (in *italics*) and the year of publication. For example, one author: (*Naatanen*, 1994); two authors: (*Kraus* and *MacGee*, 1994); three authors: (*Picton* et al., 1994). Works submitted for publication but not yet accepted are marked »Unpublished« and not included in the bibliography.
The bibliography should list only the publications quoted in the text. Names and initials (without a full stop) of authors should not be separated by commas. Only the names of different authors are separated by commas. All the authors of a publication must be listed – »et al.« is not sufficient. The letters a, b, c etc. are added to the date of publication in order to distinguish between several works published in the course of one year by the same author, e.g. (*McGee*, 1994a).
Examples
Work published in periodicals, magazines etc.:
Naatanen R, Picton T (1987) The NI wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. Psychophysiology 24, 37–425
Monographs:
Naatanen R (1992) Attention and Brain Function. Lawrence Erlbaum Associates, Hillsdale, NJ
Publications in books:
Kraus N, McGee T (1994) Auditory event-related potentials. In: Katz J (ed.) Handbook of Clinical Audiology. Williams & Wilkins, Baltimore, Hongkong, London, Munich, Sydney, Tokyo, 403–423
Curriculum vitae
All contributors are requested to submit portrait photographs (colour or black and white) and a short resume in English and German which also includes their full mailing and E-mail addresses.

Neuerscheinung



2002, 198 Seiten, broschiert,
€ 21,– / sfr 40,–
ISBN 3-922766-78-1

Aus dem Inhalt:

- Ressourcenorientierung in der Zusammenarbeit mit Familien hörgeschädigter Kinder
- Kohärenzgefühl – ein neues, möglicherweise nützliches Konzept für die Hörgeschädigtenpädagogik
- Empirische Befunde zur Bedeutung des Kohärenzgefühls bei Eltern hörgeschädigter Kinder
- Kohärenzgefühl und Praxis der Hörgeschädigtenpädagogik – Plädoyer für eine psychosoziale Umgestaltung der Beratungs- und Förderangebote

Im Zusammenhang mit kritischen Lebensereignissen und deren konstruktiver Bewältigung hat in den letzten Jahren mit dem Kohärenzgefühl von Aaron Antonovsky ein Konzept Eingang in die Sozialwissenschaften gefunden, von dem man sich innovative und klärende Impulse gleichermaßen verspricht. Somit ist es insbesondere auch für den Bereich der Behindertenpädagogik von Interesse, die Bedeutung des Kohärenzgefühls als personale Ressource (neben anderen relevanten Merkmalen wie z.B. materiellen, sozialen Ressourcen) im Copingprozess von Eltern behinderter Kinder genauer zu beschreiben. Das vorliegende Buch fasst hierzu die Ergebnisse einer empirischen Studie an 330 Eltern hörgeschädigter Kinder zusammen. Zusätzlich werden auf der Basis zahlreicher Befunde die Chancen eines ressourcenorientierten Vorgehens gegenüber einem methodenorientierten Vorgehen für die Hörgeschädigtenpädagogik aufgezeigt.

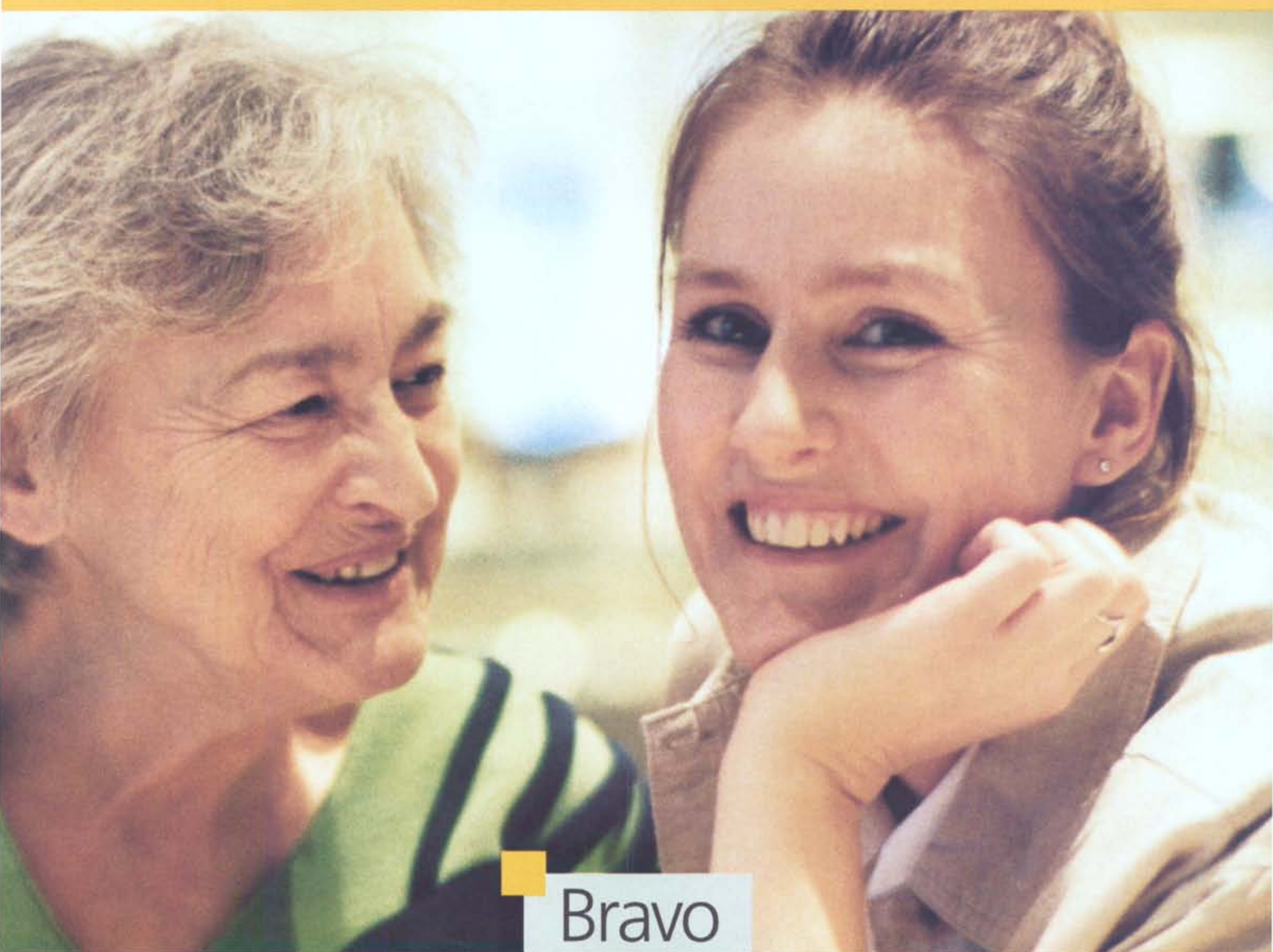
Bestellen Sie beim:

Median-Verlag von Killisch-Horn GmbH
Postfach 10 39 64 · 69029 Heidelberg
Tel. 06221/905090 · Fax 06221/9050920
E-mail: median-verlag@t-online.de

Auslieferung Schweiz:

Roggen-Amrein
AG für Kommunikation und Marketing
Rainstrasse 2a · CH-5415 Nussbaumen
Tel. 056/282 25 65 · Fax 056/282 25 33

To Analogue!



Bravo - Digital for Everyone!

Fachbücher aus dem Median-Verlag

»Zur Psychologie und Soziologie Hörbehinderter«

von Johannes Eitner

Fachwissenschaftliche Reihe der Akademie für Hörgeräte-Akustik, Band 1
2., überarbeitete Auflage 1996, 152 Seiten, Hardcover, ISBN 3-922766-25-0
€ 35,– / sfr 59,–

Inhalt:

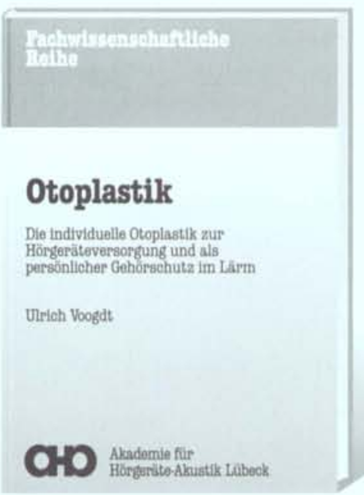
- Zur psychischen und sozialen Situation Hörgeschädigter.
  Die soziale Eingliederung Hörgeschädigter.
  Die Bedeutung des Gehörsinnes für den Menschen.
  Zum Umgang Hörender und Hörgeschädigter miteinander.
- Der ältere Mensch – die Hauptklientel des Hörgeräte-Akustikers. Entwicklung, Prognosen. Morphologische und physiologische Veränderungen im Alter. Psychologische Veränderungen im Alter.
- Verkaufs- und Beratungspsychologie. Die Bedeutung verkaufs- und beratungspsychologischer Kenntnisse für den Hörgeräte-Akustiker. Begriffserklärungen. Anpassung und Verkauf – soziale Interaktionen, Kommunikationsanalyse. Zum Einsatz der Kommunikationselemente. Kommunikationsebenen. Aktives Zuhören. Richtiges Fragen. Motivieren. Demonstrieren und Aktivieren. Der Hörgeräte-Akustiker und sein Kunde – die Stationen im Hörgerätefachgeschäft aus psychologischer Sicht.

»Otoplastik«

von Ulrich Voogdt

Die individuelle Otoplastik zur Hörgeräte-Versorgung und als persönlicher Gehörschutz im Lärm

Fachwissenschaftliche Reihe der Akademie für Hörgeräte-Akustik, Band 2
2., völlig überarbeitete Auflage 1998, 344 Seiten, Hardcover, ISBN 3-922766-28-5
€ 52,– / sfr 89,–



Inhalt:

- Die Bedeutung der Otoplastik für die Hörgeräte-Versorgung.
- Die Notwendigkeit und Problematik der individuellen Otoplastik.
- Werk- und Hilfsstoffe.
- Ohrabformung als Grundlage für ein passgenaues Ohrstück.
- Der gelungene Umweg über das Negativ oder direkt zum Vorprodukt der Otoplastik – die Rohlingherstellung.
- Aus dem Rohling wird eine Maß-Otoplastik.
- IO-Schalenfertigung.
- Kinder-Otoplastiken.
- Sonderformen wie Stütz-Otoplastik, Epithesen und Obturatoren.
- Prüfen und Messen.
- Der individuelle persönliche Gehörschutz.
- Laborausrüstung.
- Gefährliche Arbeitsstoffe.

Bestellen Sie beim:
Median-Verlag von Killisch-Horn GmbH
Postfach 10 39 64 · 69029 Heidelberg
Tel. 06221/9050915 · Fax 06221/9050920
E-mail: median-verlag@t-online.de

Auslieferung Schweiz:
Roggen-Amrein
AG für Kommunikation und Marketing
Rainstraße 2a · CH-5415 Nussbaumen
Tel. 056/282 25 65 · Fax 056/282 25 33
E-mail: rakom@pop.agri.ch

»Zeitschrift für Audiologie« am Scheideweg

Als Leserin oder Leser der »Zeitschrift für Audiologie« mögen Sie sich zu Recht gefragt haben, warum Heft 1/2002 so lange hat auf sich warten lassen. Der Grund für das verspätete Erscheinen des vorliegenden Hefts ist ebenso einfach wie beunruhigend: Leider lagen trotz aller Akquisitionsbemühungen über Monate keine druckreifen Manuskripte vor.

Und obwohl die Audiologie in den deutschsprachigen Ländern zweifellos eine ausreichend breite Basis für einen regen wissenschaftlichen Meinungsaustausch bietet, das belegt nicht zuletzt die erfreuliche Entwicklung der Jahrestagungen der Deutschen Gesellschaft für Audiologie, kommt das »Austrocknen« der »Zeitschrift für Audiologie« – so lange sie sich als rein wissenschaftliches Organ versteht – durchaus nicht überraschend. So muss eine Zeitschrift, die wissenschaftlichen Anspruch erhebt, es aber nicht schafft, in Current Contents oder anderen Indizes gelistet zu werden, zwangsläufig unter Druck geraten, weil sie ihren Autoren keinen Impact Faktor bieten kann. Diese Logik hat sich in den letzten Jahren in einem alarmierenden Rückgang des Manuskriptaufkommens niedergeschlagen, da potentielle Autoren im Hinblick auf die Karriereplanung bei ihren Publikationsüberlegungen heute vorrangig – anders als noch vor Jahren – auf den Impact Faktor achten müssen. Da es die »Zeitschrift für Audiologie« bisher leider nicht geschafft hat, einen Impact Faktor zu bekommen, und dies auch mittelfristig nicht zu erwarten ist, hat sich die Situation derartig zugespitzt.

Angesichts dessen wird man sich über die weitere Zukunft der »Zeitschrift für Audiologie« Gedanken, und leider auch Sorgen machen müssen. Wie könnte die Zukunft unserer Zeitschrift gestaltet werden und welche Optionen bieten sich an?

1. Im günstigsten Falle gelingt es, das Publikationspotenzial derer zu erschließen, die nicht (mehr) in so starkem Maße dem karrierebedingten Publikationsdruck unterliegen – also verstärkt etablierte Kolleginnen und Kollegen als Autoren zu gewinnen. Wenn man das erreichen kann, mag es mittelfristig auch gelingen, in die Indizes aufgenommen zu werden. Damit wäre die Zeitschrift für eine breite Autorenschaft als Publikationsforum interessant und der Fortbestand der Zeitschrift als wissenschaftliches Journal wäre gesichert. Ob allerdings eine *ausreichende Zahl von Autoren bereit ist, dieses Opfer – also freiwilligen Verzicht auf Publikationsmöglichkeiten in angesehenen Zeitschriften mit internationaler Verbreitung – über einen ausreichend langen Zeitraum zu erbringen, muss ernsthaft in Frage gestellt werden.*

2. Ferner wäre eine Fusion mit einer oder mehreren bestehenden Zeitschriften in Erwägung zu ziehen. Damit könnte die »kritische Masse« überschritten werden und es könnten Synergieeffekte zum Vorteil aller beteiligten Organe ausgenutzt werden. Über mögliche Fusionspartner kann man derzeit nur spekulieren. Zunächst müsste ein überzeugendes diesbezügliches Konzept entwickelt werden, bevor man mit potenziellen Partnern in konkrete Verhandlungen eintreten kann. Immerhin existiert seit Beginn diesen Jahres ein prominentes, internationales Exempel für ein derartiges Vorgehen: das »International Journal of Audiology« als Fusionsprodukt der Zeitschriften »Audiology«, »British Journal of Audiology« und »Scandinavian Audiology«. Dieses Beispiel verdeutlicht, dass es auch auf internationaler Ebene offenbar schwierig ist, eine audiologische Fachzeitschrift erfolgreich auf Kurs zu halten.

3. Die dritte, und wohl realistischste Option wäre eine Umwandlung der Zeitschrift vom Publikationsorgan mit vorrangig wissenschaftlichem Anspruch in ein Mitteilungsblatt der Deutschen Gesellschaft für Audiologie (DGA), das neben Originalmitteilungen und Übersichtsartikeln verstärkt auch Fortbildungsartikel, Kongressberichte, Mitteilungen und Berichte aus verschiedenen Arbeitsgruppen und Forschungslaboren veröffentlicht. Dieser Ansatz bedarf erfahrungsgemäß einer intensiven Einwerbung von Artikeln wie auch der Bereitschaft auf der Autorenseite, Vortragsmanuskripte etc. für diesen Zweck zu bearbeiten und Fachartikel auf Einladung bzw. Aufforderung zu verfassen.

Neben diesen drei Entwicklungsmöglichkeiten mögen durchaus noch andere Entwicklungsstrategien in Frage kommen, und es wäre interessant, aus der Leserschaft diesbezügliche Anregungen zu erhalten.

Was also kann konkret unternommen werden? Auf Anregung des Präsidiums der Deutschen Gesellschaft für Audiologie (DGA) werden die Herausgeber die Initiative im Sinne der dritten Option ergreifen und verstärkt Fortbildungs- und andere Fachartikel einwerben. Daneben bietet sich die Zeitschrift für Audiologie selbstverständlich wie bisher für die Veröffentlichung von Originalien und Übersichtsartikeln an, um ihrer Leserschaft eine interessante Mixtur von Beiträgen bieten zu können. Dies erfordert die aktive Mitwirkung möglichst vieler Autorinnen und Autoren aus allen Bereichen der Audiologie im Geist der Satzung der DGA.

Das bedeutet also: Entweder ein bisschen mehr Arbeit für alle oder, wenn Sie sich als Autor(in) nicht in stärkerem Maße beteiligen können oder wollen, das baldige Ende der »Zeitschrift für Audiologie«.

Fragen Sie also nicht, in Abwandlung eines bekannten Kennedy-Zitats, was die Zeitschrift für Sie leisten kann, sondern fragen Sie sich selbst, was Sie für unsere Zeitschrift tun können. Dann brauchen wir um die Zukunft der »Zeitschrift für Audiologie« nicht länger besorgt sein!

Jürgen Kießling

Noise reduction schemes for digital hearing aids – Part I: Listening effort and speech intelligibility

Mark Marzinzik, Birger Kollmeier
AG Medizinische Physik, Universität Oldenburg, D-26111 Oldenburg

Abstract *Subjective methods for the evaluation of benefits from different noise reduction schemes are proposed and tested with three different single-microphone algorithms (Ephraim and Malah 1984, 1985) and a binaural directional filter and dereverberation algorithm (Wittkop 2001). In this first contribution, listening effort assessed by a newly developed procedure and speech intelligibility are considered. The measurements were carried out in a sound-insulated booth with six normal-hearing and six moderately sensorineural hearing-impaired subjects. With respect to speech intelligibility, the Ephraim-Malah single-microphone noise reduction algorithms revealed no improvement over the unprocessed signal. However, benefits with respect to reductions in listening effort were found when the listening effort test was carried out. The binaural algorithm showed slight but significant improvements in speech reception thresholds (SRT). In addition, listening effort benefited to some extent in speech-shaped noise with this algorithm. The results indicate that conventional SRT tests and tests of listening effort appear to measure different aspects of the effect of noise reduction schemes in speech perception. Also, binaural noise suppression appears to be more effective in enhancing speech intelligibility than single-microphone algorithms.*

Key words: *noise reduction*
speech enhancement
listening effort
speech intelligibility
digital hearing aids

Corresponding authors: Dr. Mark Marzinzik, Prof. Dr. Dr. Birger Kollmeier
AG Medizinische Physik
Carl von Ossietzky Universität Oldenburg
D-26111 Oldenburg
Phone +49 (0)441 7985466, Fax +49 (0)441 7985466
E-Mail:mark@medi.physik.uni-oldenburg.de, Birger.Kollmeier@uni-oldenburg.de

Störgeräuschunterdrückungsalgorithmen für digitale Hörgeräte – Teil I: Zuhöranstrengung und Sprachverständlichkeit

Mark Marzinzik, Birger Kollmeier
AG Medizinische Physik, Universität Oldenburg, D-26111 Oldenburg

Zusammenfassung *In dieser Arbeit werden die monauralen Störgeräuschunterdrückungsalgorithmen von Ephraim und Malah (1984, 1985), die in Marzinzik und Kollmeier (2001) detailliert vorgestellt wurden, sowie ein binauraler Richtungsfilterungs- und Enthallungsalgorithmus von Wittkop (2001) hinsichtlich ihrer Sprachverständlichkeit sowie ihrer Zuhöranstrengung untersucht.*

Die mentale Anstrengung, die nötig ist, einem Sprecher in stark störgeräuschbehafteter Umgebung zuzuhören, wird dabei mit einem neu entwickelten Testverfahren ermittelt. Das vorgeschlagene Verfahren beinhaltet eine anstrengende Zuhöraufgabe, die sicherstellen soll, dass tatsächlich die Zuhöranstrengung erfasst wird und nicht lediglich bessere Klangqualität.

Die Experimente wurden mit sechs normalhörenden sowie mit sechs sensorineural schwerhörigen Versuchspersonen in einer schallisolierten Hörkabine durchgeführt. Obwohl eine starke Störgeräuschunterdrückung mit den monauralen Algorithmen erreicht wird, schlägt sich dieses nicht in verbesserter Sprachverständlichkeit nieder, wie sie in der vorliegenden Arbeit mit Satztests ermittelt wurde. Es besteht im Gegenteil eher eine Tendenz zur Verringerung der Verständlichkeit. Mit dem binauralen Störgeräuschunterdrückungsalgorithmus von Wittkop konnten dagegen geringe Verbesserungen der Sprachverständlichkeit nachgewiesen werden.

In Bezug auf die Zuhöranstrengung konnte eine signifikante Verbesserung mit einem der drei Ephraim-Malah-Algorithmen (EL) unter Bohrmaschinenstörgeräusch festgestellt werden. In fluktuierendem Cafeteria-Störgeräusch wurden allerdings keine signifikanten Effekte gefunden.

Hinsichtlich der Zuhöranstrengung wurden Unterschiede zwischen Algorithmen gefunden, die sich nicht in den Ergebnissen der Sprachverständlichkeitsmessungen abbilden. Dies kann als ein Hinweis darauf verstanden werden, dass der entwickelte Test auf Zuhöranstrengung tatsächlich andere Aspekte der Störgeräuschunterdrückung erfasst als konventionelle Sprachverständlichkeitstests.

Schlüsselwörter: *Störgeräuschunterdrückung*
Sprachverbesserung
Zuhöranstrengung
Sprachverständlichkeit
digitale Hörgeräte

1. Introduction

Noise evokes major listening difficulties in hearing-impaired subjects, even in persons with low to moderate hearing losses (*Weiss* and *Neuman* 1993). These difficulties are often experienced as a real handicap, especially at the working place and during social activities. They are connected with decreased speech intelligibility and with an enormously increased effort to understand speech in noise. Noise reduction schemes in digital hearing aids may help to overcome these deficiencies by increasing the signal-to-noise ratio. They aim at reducing speech reception thresholds, i.e. by increasing speech intelligibility, lowering the listening effort and improving the perceived quality of the acoustic environment.

However, the literature on noise reduction indicates that single-microphone noise reduction brings little or no improvement to speech intelligibility if a speech signal is degraded by wideband noise (for a review, see the book from *Studebaker* and *Hochberg* 1993; cf. also *Marzinzik* and *Kollmeier* 2001). On the other hand, listening to the algorithms confirms that the noise *is* actually reduced in general. One potential benefit of such algorithms is supposed to be increased »ease of listening« (less listening effort). Indeed, fatigue and increased effort when listening in noise is a common complaint of hearing-impaired subjects. Even normal-hearing subjects suffer from this complaint after a day of exposure to noise at work. However, speech recognition tests did not reflect this fatigue (*Ivarsson* and *Arlinger* 1993), so it may well be related to non-auditory functions, involving concentration, attention et cetera.

Figure 1 shows an optical illustration of noise effects. A short discussion of it may help us to understand it acoustically. Figure 1a gives an example of how noise disturbs a target signal. The text is blurred and noise suppression tries to bring it back into focus. Most probably, a reading test with subjects will not reveal any significant differences in the error rate between the blurred and noise-reduced text, as the subjects are able to counteract the blurring of the text by concentrating harder. However, after having read a long text, the subjects with the blurred text are generally more tired than those who have read a clean text. But how can this fatigue be assessed? If the subjects have to perform a reading test after a preceding »fatiguing« phase, they will yet again most probably be able to counteract the fatigue by increased concentration. As most subjects will try not to disappoint the person conducting the experiment, it can be expected that they will do their very best to overcome the fatigue. This is known as the Hawthorne effect (*Roethlisberger* and *Dickson* 1939). Hence, a conventional reading test (and in the acoustics field, this would mean a conventional speech intelligibility test) seems to being not adequate when it comes to assessing the fatiguing effect of noise. A possible solution is to let the subjects read two long texts (blurred and clean) and have them assess their own fatigue (or, alternatively, the effort that was needed or the »ease of reading«). This procedure led to the listening effort test presented in this paper.

A different approach is based on the conjecture that response times are longer when the signal is noisy rather than clean. In the case of the noise in the optical example of Figure 1a, however, it seems more likely that any presumed increase in the reading time is too small to be reliably detected and might perhaps even be compensated by heightened concentration. On the other hand, if the signal-to-noise ratio deteriorates so much that whole signal parts actually get lost, longer response times are probable. The reader can check this in Figure 1b.

Gatehouse (1994) suggested a »sentence verification test« in which response times are measured to assess listening effort. This test was applied by *Baer* et al. (1993) to evaluate a spectral contrast enhancement technique. They found effects for the response times twice as many as for the intelligibility scores and they were statistically more robust. Measuring response times in an intelligibility test had already been proposed by *Hecker* et al. (1966). They suggested that response times provide an independent measure and can increase the sensitivity of conventional tests. Another approach was applied by *Hoeks* and *Levelt* (1993). They used pupillary dilation as a measure of the level of mental effort needed in an attentional task. However, to obtain stable (and reliable) results, the measurement of response times as well as the measurement of pupillary dilation need averaging over numerous trials since the trial-to-trial variability is large.

Dillon and *Lovegrove* (1993) (referring back to *Lim* and *Oppenheim* 1979) point out that long-term listening at a reduced fatigue level may lead to a long-term gain in intelligibility if the listener tires more quickly without a noise reduction system. This connects listening effort with speech intelligibility on a larger time scale. If this holds true, listening effort measurements might be used to estimate long-term speech intelligibility.

For these reasons, we are introducing a new procedure here, which is based on subjective self-assessment. This procedure is supposed to be sensitive to the effects of noise on speech as regards listening effort at arbitrary signal-to-noise ratios.

In addition, conventional speech reception threshold (SRT) tests were also employed to compare listening effort and speech intelligibility measurements for the same subjects and the same set of algorithms and test signals.

2. Method

Speech intelligibility and listening effort were measured using two types of noise reduction algorithms.

Fig. 1: Optical illustration of noise effects.

Abb. 1: Optische Illustration von Störgeräuscheffekten.

Single-microphone noise reduction algorithms were chosen since they yield large reductions in different (stationary) background noises. However, these schemes are only capable of reducing noises which do not change too drastically between two speech pauses (in which noise estimates might be updated).

Multi-microphone algorithms are the only ones that may provide improvements in speech intelligibility over a wide range of noises (wide-band especially) as is indicated by the literature on noise reduction reviewed in *Marzinzik* and *Kollmeier* (2001).

2.1 Selected algorithms

The single-microphone noise reduction schemes proposed by *Ephraim* and *Malah* (1984, 1985) overcome the annoying »muscial tones« artifact of conventional schemes based on spectral subtraction, while keeping computational complexity relatively low. Three different versions were employed in Experiment 1: The minimum mean-square error (MMSE) short-time spectral amplitude (STSA) estimator (Eq. 7 in *Ephraim* and *Malah* 1984; called E7 in the following), a modified estimator under uncertainty of signal presence (Eq. 30 in *Ephraim* and *Malah* 1984; called E30 in the following), and the MMSE log-spectra estimator (*Ephraim* and *Malah* 1985; called EL in the following). The implementations of these algorithms use the decision-directed approach for estimating the a priori signal-to-noise ratio. An important prerequisite for the application of these algorithms in hearing aids is their combination with an automatic procedure that updates the noise spectrum estimate, since the acoustic environment is supposed to change over time. For this reason, we used the speech pause detection algorithm proposed in *Marzinzik* and *Kollmeier* (2002b) with low computational complexity and a relatively low false-alarm rate. The noise spectrum estimate is updated during detected speech pauses.

In Experiment 2, a binaural noise reduction algorithm was applied. The two-microphone directional filter and dereverberation algorithm developed by *Peissig* (1993) was shown to yield significant improvements in speech intelligibility (*Kollmeier* et al. 1993). The directional filtering stage of this algorithm attenuates lateral sound sources while passing through sounds from the front. The dereverberation stage reduces diffused noise and reverberation. This algorithm was further elaborated and improved upon by *Wittkop* et al. (1999) and *Wittkop* (2001). The main objective was to preserve a high signal quality in the processed signal. This algorithm is called DD in the following. Moreover, single-microphone algorithm E7 as well as the sequential processing of the binaural algorithm DD and the single-microphone algorithm E7 (denoted as DDE7) were applied in Experiment 2. While the processing by DD is assumed to reduce reverberation and diffused noise as well as distinct noise sources separated in space, the additional processing by E7 in algorithm DDE7 is assumed to further reduce stationary components of the background noise.

2.2 Subjects

Noise reduction is expected to yield benefits even for normal-hearing listeners, not only for hearing-impaired subjects. However, an evaluation using only normal-hearing listeners may be misleading because the results from normal-hearing subjects can differ significantly from those of hearing-impaired subjects. *Levitt* et al. (1993), for example, report that normal-hearing subjects showed a significant decrement while hearing-impaired subjects showed a significant improvement in consonant recognition using a Wiener filtering noise reduction algorithm. *Hygge* et al. (1992) found that hearing-impaired and normal-hearing persons differ substantially in how they are affected by background noise when trying to comprehend foreground speech. For this reason, all measurements in this paper were primarily carried out on hearing-impaired subjects. To investigate whether any systematic differences existed between normal-hearing and hearing-impaired listeners with respect to noise reduction processing, Experiment 1 was also performed using six normal-hearing listeners.

The six normal-hearing subjects (three male and three female students aging from 21 to 29 years) had no prior experience of psychoacoustic measurements, no history of hearing problems and pure tone thresholds of less than 10 dB HL for at least seven of the nine audiometric frequencies between 125 Hz and 8 kHz. In addition to these, six subjects with moderately bilateral sensorineural hearing losses (three males and three females aging from 23 to 78 years) participated in both the evaluation of the single microphone as well as the binaural noise reduction algorithms. Figure 2 gives the median and the range (minimum to

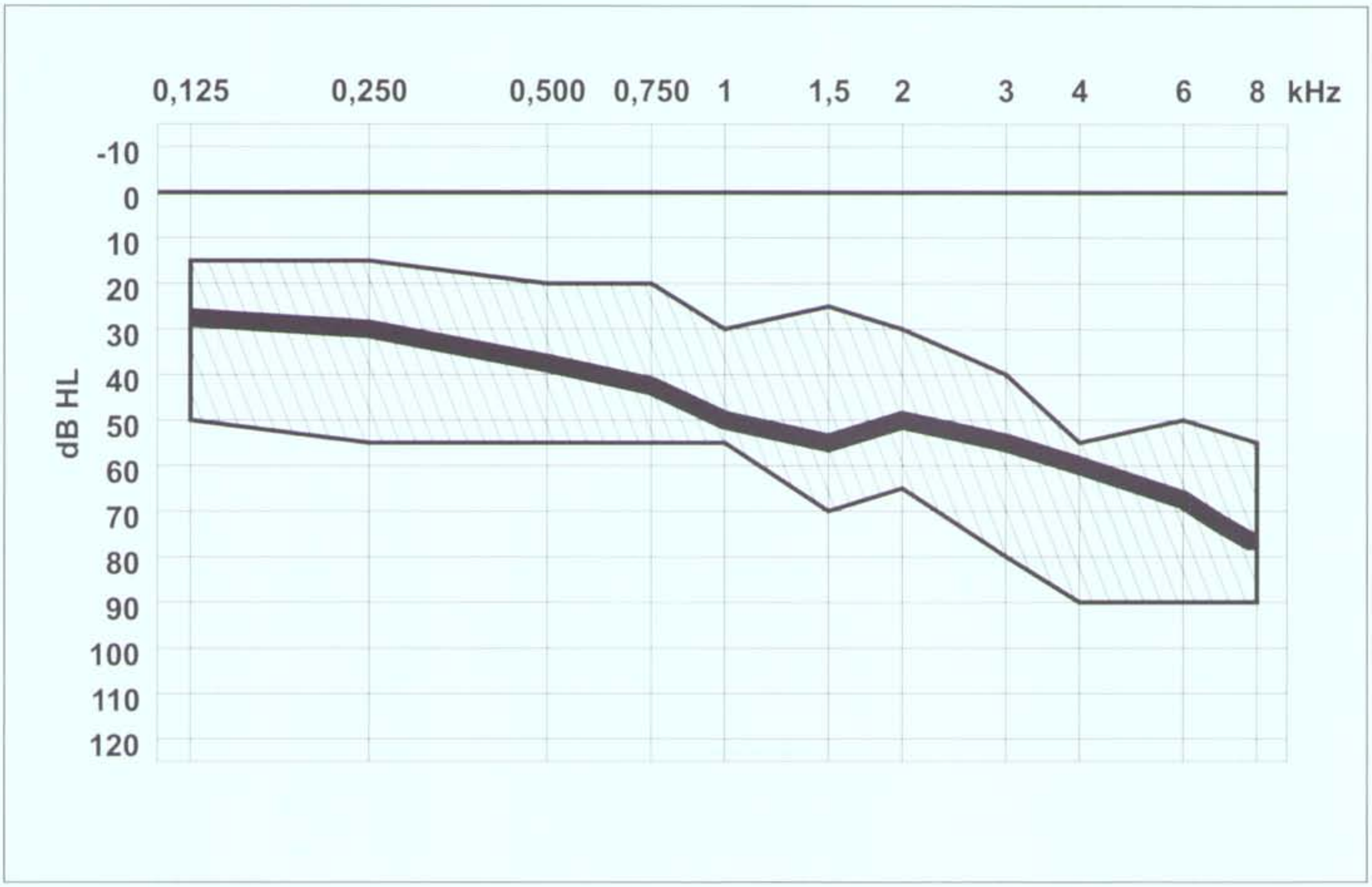


Fig. 2: Median audiogram of the hearing-impaired subjects. The solid line indicates the median hearing loss of the left and right ears of the six hearing-impaired subjects. The shaded area shows the total range.

Abb. 2: Audiogramm der schwerhörigen Versuchspersonen. Die durchgezogene dicke Kurve gibt den Median des Hörverlustes beider Ohren der sechs Versuchspersonen an. Die schattierte Fläche zeigt den gesamten vorkommenden Bereich an.

maximum) of the hearing losses. All were experienced listeners who had been involved in other psychoacoustic measurements.

2.3 Measurement setup

The subjects were seated in a sound-insulated booth. All experiments were computer-controlled. In the listening effort measurements, the subjects were allowed to switch between algorithms using a handheld touchscreen response box. In the speech intelligibility measurements, the person conducting the experiment was also seated in the booth and operated the computer by using the touchscreen response box.

To approximately compensate for the hearing loss of the hearing-impaired subjects at intermediate levels, third-octave band equalization (a range of ± 16 dB in each band) was employed for each ear independently (28-band graphic equalizer). The required gain in each frequency band was determined from the audiogram of each subject using the one-half gain rule, i.e. target gain at each frequency is simply the subject's audiometric threshold multiplied by 0.5 (*Lybarger* 1944, 1978). Most threshold-based approaches in use today are modifications of this simple one-half gain rule (*Fabry* and *Schum* 1994). The actual frequency response of the equalizers was verified by a spectrum analyzer.

In the listening effort measurements, the unprocessed signal (i.e., without noise reduction) and the signals processed by the three noise reduction algorithms under consideration (cf. Table 1) were recorded synchronously on four tracks of a digital audio tape recorder. During playback, the output was amplified and presented to the subjects via headphones (Sennheiser HD 25).

In the speech intelligibility measurements, the additional amplification was provided by an audiometric amplifier and a Sennheiser HDA 200 headphone was used.

Table 1: Noise signals used in the measurements.

Tabelle 1: In den Messungen benutzte Störgeräuschsignale.

Noise signal		Listening effort		Speech intelligibility	
		Experiment 1 (UN,E7,EL,E30)	Experiment 2 (UN,E7,DD,DDE7)	Experiment 1 (UN,E7,EL,E30)	Experiment 2 (UN,E7,DD,DDE7)
Monaural recordings					
Drill	Drilling machine	× (-5 dB SNR)		×	
Caf_m	Cafeteria noise	× (0 dB SNR)		×†	
Binaural recordings					
Ssn_{60}	Speech-shaped noise (from 60°)		× (5 dB SNR)		×
Caf_b	Cafeteria noise (diffuse)		× (5 dB SNR)		

† Only UN and the best noise reduction algorithm from the drill noise measurement

Table 1 gives an overview of the noise signals used in the different measurements.

In each experiment, both a stationary and a fluctuating noise were employed. In Experiment 1, noise samples recorded from a drilling machine were used, since these represent a class of technical equipment noises which are often rather stationary. In addition, performance in cafeteria noise with babble in the background was tested. This noise fluctuates greatly and represents a noisy environment typical of social gatherings. In Experiment 2, binaural recordings were applied. The target speech signals and the speech-shaped noise signal were recorded with a Head Acoustics dummy head in a seminar room using a reverberation time of T_{60} = 0.6 s. The speech-shaped noise was presented from a 60° incidence direction and at 1 m distance from the dummy head. The target speech was always presented from the front. The signal-to-noise ratios were set by measuring long-term RMS values (over the whole length of the speech sample) in the right channel and attenuating or amplifying the speech signal accordingly. This channel of the binaurally processed signals was presented diotically to the subjects via headphones. Hence, any binaural cues that could be used by the subjects' own binaural processing system were eliminated. Since binaural capabilities generally vary a great deal among hearing-impaired subjects, this was necessary when using only six subjects. The presentation level was individually adjusted so that perception was »loud but still comfortable« to guarantee that most signal parts were audible to the subject. No freefield correction was applied.

2.4 Listening effort test

Different local radio newscasts that were at least two years old had been re-recorded in a radio studio, spoken by a professional male newscaster. The newscasts were put together to give blocks of 2 1/2 minutes and were mixed afterwards with different noises at different signal-to-noise ratios (see Table 1). The noisy newscasts were then processed by the noise reduction algorithms under consideration.

During the measurement session, part of a radio newscast is first presented to the subject (seated in a sound-insulated booth) in order to adjust the overall gain so that the overall loudness impression is at the top border of the comfortable loudness range. Thereafter, the subject is requested to listen to a newscast of 2.5 min duration, mixed with background noise. The task is to listen carefully to the news and afterwards repeat as much news as can be remembered (Step 0). What the subjects repeats is recorded on a dictaphone by the person conducting the experiment. However, the content of the subject's report of the newscast is not evaluated. According to *Jones* (1983), cognition experiments

Table 3: Listening effort scale as recommended by the International Telecommunication Union (ITU 1996).

Tabelle 3: Von der ITU empfohlene Skala der Zuhöranstrengung (ITU 1996).

Effort required to understand the news	Score
Complete relaxation possible; no effort required	5
Attention necessary; no appreciable effort required	4
Moderate effort required	3
Considerable effort required	2
No meaning understood with any feasible effort	1

revealed that the overall number of words reported is not reduced in noise, but that they are produced in a fashion which is both less coherent and organized. In this test, however, the repetition procedure only serves to put the subjects under stress and to force them to really listen carefully (thereby probably producing an additional fatigue effect). This was the reason for choosing news that was at least two years old and only locally relevant, and thus unknown to the subjects. Step 0 is introduced so that the subject can practise the listening task and become aware of the distraction caused by the background noise.

In the next step (Step 1), the subject has to listen to the same newscast again, but now has the opportunity of switching between the different programs. This is done by the subject on a handheld touchscreen response box with four »buttons« for switching between the different algorithms. The subject is instructed to try all four programs and to judge them according to the 5-point listening effort scale recommended by the *ITU* (1996)[1] (Table 3). This scale is listed on an assessment form which the subject is required to fill in after listening. The numerical scores are not visible to the subject. They are only relevant for the evaluation of the responses.

In addition, the subject has the opportunity of commenting on the different programs orally or in writing. Step 1 was introduced to let the subject get accustomed to the available algorithms and to try them out while not under pressure.

In the following step (Step 2), the subject is requested to listen to a second unknown newscast and to fill in a new assessment form for the programs (the person conducting the experiment removes the first one). The task is to again repeat as much news as can be remembered. This task is meant to put pressure on the subject in order to really assess the listening effort associated with the algorithms instead of primarily judging the sound quality. The instructions are slightly different than before: The subject can switch between the four programs at will, but is also allowed to stay in a program if listening is easier with that particular one. After repeating the newscast, the subject fills in the assessment form again.

Depending on the question under study, the procedure used in Step 2 can be repeated yet again with a different noise and/or signal-to-noise ratio and/or speaker. In the present studies, a female newscaster and cafeteria noise were chosen as Step 3.

2.5 Speech intelligibility test

In Experiment 1, the Göttingen sentence test (*Kollmeier* and *Wesselkamp* 1997) was applied to determine the speech reception thresholds (SRT) for the different algorithms, i.e. the signal-to-noise ratio at which 50 % of the speech is correctly understood. The sentences of the Göttingen sentence test, which were spoken by an unschooled male speaker, are combined in lists of 10 sentences each. In Experiment 1, one test list served to determine speech intelligibility at one fixed signal-to-noise ratio (SNR). For a reliable estimation of the psychometric function (speech intelligibility versus SNR) three points of this function were determined individually for each condition by measuring at three different SNRs. All test sentences were mixed with drilling machine noise and cafeteria noise, respectively, at SNRs ranging from -26 to +16 dB and were processed offline with the noise reduction algorithms under consideration. The overall gain was adjusted in a preliminary test run for each subject so that the overall loudness impression was at the top end of the comfortable loudness range. The subjects were seated in a sound-proofed booth together with the person conducting the experiment and given the task of repeating what they heard. The correct sentence is displayed to the person conducting the experiment on a handheld touchscreen response box and he or she then marks the words that were either not heard or repeated incorrect by the subject. Then the next sentence is presented. Table 1 shows the conditions that were tested. Both the order of the test lists and the order of the algorithms were chosen at random. A logistic function was fitted to the measured data using a maximum likelihood method. In this way, it was possible to determine the speech reception threshold (SRT) and the slope of the psychometric function.

Recently, another German sentence test was developed – the Oldenburg sentence test (*Wagener* et al. 1998, 1999). Because

[1] The following German translations were used for the measurements:
Völlige Entspannung ist möglich, keine Anstrengung erforderlich (5); Aufmerksamkeit ist erforderlich, aber keine nennenswerte Anstrengung nötig (4); Mäßige Anstrengung ist erforderlich (3); Beträchtliche Anstrengung ist erforderlich (2); Trotz größter Anstrengung ist die Bedeutung unverständlich (1).

the limited set of sentences in the Göttingen sentence test restricts the possible number of conditions that can be measured without learning effect, the Oldenburg sentence test was applied in Experiment 2. It uses sentences that are syntactically correct, but make no sense. The sentences (as well as the test lists) can be used repeatedly during the measurement because of their nonsense character. The test lists used consisted of 20 sentences each.

2.6 Statistical methods

Our report will include the individual results as well as the median values over all subjects and the median absolute deviations (MAD). The latter is used to describe variation. Its advantage over the ordinary range or standard deviation is that it is not sensitive to extreme outliers. Furthermore, it does not assume a Gaussian distribution of the data like the standard deviation. Compared to the interquartile range (IQR), the MAD is even more robust. It is defined as

$$\mathrm{MAD}(x_1, \ldots, x_N) = 1.4826 \cdot \mathrm{median}(\,|x_1 - \mathrm{median}(x_1, \ldots, x_N)|, \ldots, |x_N - \mathrm{median}(x_1, \ldots, x_N)|\,), \qquad (1)$$

where $x_1, \ldots, x_N$ is the respective data set. The constant 1.4826 ensures that the MAD approximates the standard deviation σ if the data has a Gaussian bell-shaped distribution.

A Friedman two-way analysis of variance by ranks test is applied to the experimental results to find out whether there are any significant differences between the algorithms. In all experiments, a difference is only regarded as significant if the P value of the Friedman chi-square statistics (χ_r^2 with df degrees of freedom) is below $\alpha = 0.05$, otherwise it is regarded as not significant. According to *Motulsky* (1995), many statisticians avoid terms like »very significant« or »extremely significant« and think that the word »significant« should never be prefaced by an adjective, since once an α level is set, a result is simply statistically significant or not.

One method of testing among the subjects' ratings for the overall concordance is Kendall's W coefficient of concordance. Kendall's W is a normalization of the Friedman statistic, so the χ^2 and df values are the same. Table 2 gives a classification of the W values.

If the Friedman test indicates significant differences between algorithms, it is worth taking a closer look at the data. Dunn's post test for multiple comparisons can be used for this. In general, Wilcoxon's matched pairs signed rank test is used to find out which algorithms differ significantly from each other. However, it is not appropriate to repeatedly use a Wilcoxon test for the same significance level as if only two algorithms were being tested. To compare various pairs of algorithms, a correction for multiple comparisons has to be applied (*Wright* 1992). Since most corrections are too conservative, Dunn's post test for multiple comparisons is performed here instead of a Wilcoxon test.

Table 2: Classification of Kendall's W *coefficient of concordance. Although these divisions are clearly arbitrary, they do provide useful »benchmarks« for the discussion of concordance among subjects. This subdivision is adapted from a classification of the Kappa statistic by Landis and Koch (1977).*

Tabelle 2: Klassifikation von Kendalls Konkordanzkoeffizienten W. *Die Unterteilungen sind gewissermaßen beliebig, geben aber gute Anhaltspunkte für die Bewertung der Konkordanz zwischen Versuchspersonen. Die Unterteilung wurde übertragen von derjenigen der Kappa-Statistik durch Landis and Koch (1977).*

W Statistic	Strength of Concordance
0.00–0.17	Poor
0.18–0.33	Slight
0.34–0.50	Fair
0.51–0.67	Moderate
0.68–0.83	Substantial
0.84–1.00	Almost Perfect

3. Results

3.1 Listening effort

3.1.1 Experiment 1

The normal-hearing subjects' results in the listening effort test (Table 4) show a median improvement of one point with algorithm EL in drill noise compared to no noise reduction (UN) in Steps 1 and 2, while an improvement resulting from algorithm E7 is only found in Step 1, with no difference between E30 and UN. In cafeteria noise, all three noise reduction algorithms were judged worse than UN on average. The difference between UN and E7, however, is very slight (0.5 point).

The hearing-impaired subjects seemed to benefit more from the noise reduction processing than the normal-hearing subjects.

A glance at Table 5 reveals that every hearing-impaired subject reported less listening effort with at least one of the noise reduction algorithms, compared to no noise reduction in drill noise. However, there is little concordance between the subjects' judgments of the different Ephraim-Malah algorithms. While three subjects experienced the least listening effort with algorithm EL in Step 2, one subject decides on algorithm E30, another subject on algorithm E7, and yet another could not decide between algorithms E7 and EL.

Table 4: Listening effort test results of Experiment 1 with the normal-hearing subjects. Scores according to the listening-effort scale are shown for the UN (unprocessed), E7, EL, and E30 algorithms, respectively. Also given are the median values and the median absolute deviations (MAD). A high score corresponds with low listening effort. Noise conditions were drilling machine noise at -5 dB SNR and cafeteria noise at 0 dB SNR. »Without repetition« indicates that subjects didn't have to repeat the news after listening. »With repetition« means that subjects had to repeat the remembered news.

Tabelle 4: Ergebnisse der Zuhöranstrengungsmessungen mit den normalhörenden Versuchspersonen in Experiment 1. Angegeben sind die Bewertungen sowie Mediane und Medianabsolutabweichungen (MAD) für das unverarbeitete Signal (UN) sowie die Algorithmen E7, EL und E30. Höhere Werte bedeuten geringere Zuhöranstrengung. Es wurden ein Bohrmaschinengeräusch bei -5 dB SNR und ein Cafeteria-Geräusch bei 0 dB SNR untersucht. »Without repetition« deutet an, dass in diesem Fall die Versuchspersonen die Nachrichten nicht wiederholen mussten. Bei »With repetition« mussten die Nachrichten aus der Erinnerung wiederholt werden.

Normal-Hearing Subject	Step 1: Drill Without Repetition				Step 2: Drill With Repetition				Step 3: Cafeteria With Repetition			
	UN	E7	EL	E30	UN	E7	EL	E30	UN	E7	EL	E30
AA	2	4	4	4	2	3	4	3	2	4	4	4
FJ	3	3	3	3	3	3	3	3	4	4	2	2
GI	3	4	4	3	4	3	3	2	4	2	2	3
MI	5	5	4	3	4	4	4	3	4	3	2	1
MS	4	4	3	3	2	3	4	3	3	3	2	2
RE	3	4	4	4	3	3	4	3	3	3	2	2
Median	3.0	4.0	4.0	3.0	3.0	3.0	4.0	3.0	3.5	3.0	2.0	2.0
MAD	0.7	0.0	0.0	0.0	1.5	0.0	0.0	0.0	0.7	0.7	0.0	0.7

The concordance between subjects is generally lower in Step 3 (cafeteria noise) than in Steps 1 and 2. Half of the subjects benefit from the noise reduction, while the other half reported less listening effort without it.

For the normal-hearing subjects in Step 1, the results of the Friedman test were $\chi_r^2 = 4.297$, $df = 3$, $P = 0.231$ and $W = 0.239$. This indicates no significant differences between algorithms and a slight agreement among the subjects. For Step 2: $\chi_r^2 = 6.077$, $df = 3$, $P = 0.108$ and $W = 0.338$ (no significant differences, but fair agreement among the subjects); the results for Step 3 ($\chi_r^2 = 7.041$, df = 3, $P = 0.071$ and $W = 0.391$).

The results for the hearing-impaired subjects, however, show significant differences between algorithms. For Step 1, the Friedman test resulted in $\chi_r^2 = 11.659$, $df = 3$, $P = 0.009$ (moderate concordance among subjects, $W = 0.648$). Dunn's post test revealed that the improvement in listening effort using the Ephraim-Malah algorithm EL was statistically significant compared to the unprocessed signal UN. The median judgment using the unprocessed signal was »considerable effort required«, whereas with the EL algorithm the judgment on average was »complete relaxation is possible, no effort required«. Other differences were not found to be significant. For Step 2: $\chi_r^2 = 11.089$, $df = 3$, $P = 0.011$ (moderate concordance among subjects, $W = 0.616$). According to Dunn's post test, yet again the only significant difference in listening effort is between UN and EL. For Step 3: $\chi_r^2 = 7.788$, $df = 3$, approximate $P = 0.051$, exact $P < 0.05$.[2] Dunn's post test shows that, according to the rank sum differences, only the difference between E7 and E30 can be significant, i.e. E30 performs significantly worse in the cafeteria than E7 as regards listening effort. The concordance among subjects is fair ($W = 0.433$).

The statistical analysis can be summarized as follows: There were no significant differences between the algorithms as far as the normal-hearing subjects were concerned. In the case of the hearing-impaired subjects, algorithm EL was significantly better than no noise reduction (UN) with respect to listening effort in drill noise, and algorithm E30 significantly worse than E7 in cafeteria noise.

[2] The exact P values for the Friedman test can, for example, be found in Table A–22 by *Marascuilo* and *McSweeney* (1977).

Table 5: Listening effort test results of Experiment 1 with the hearing-impaired subjects.

Tabelle 5: Ergebnisse des Zuhöranstrengungstests mit den schwerhörigen Versuchspersonen in Experiment 1.

Hearing-Impaired Subject	Step 1: Drill Without Repetition				Step 2: Drill With Repetition				Step 3: Cafeteria With Repetition			
	UN	E7	EL	E30	UN	E7	EL	E30	UN	E7	EL	E30
BD	2	4	4	4	2	4	4	3	2	1	1	4
GM	1	4	5	3	1	3	5	2	5	4	1	2
HM	1	3	5	4	1	3	5	4	1	3	2	1
KF	3	4	5	5	3	3	4	5	4	3	2	2
KR	5	5	5	5	3	4	5	3	2	3	3	1
WH	2	4	5	2	3	5	4	3	2	4	2	1
Median	2.0	4.0	5.0	4.0	2.5	3.5	4.5	3.0	2.0	3.0	2.0	1.0
MAD	1.5	0.0	0.0	1.5	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.0

Table 6: Listening effort test results of Experiment 2 with the hearing-impaired subjects using algorithms UN (unprocessed), E7, DD, and DDE7. Noise conditions were speech-shaped noise and cafeteria noise at 5 dB SNR.

Tabelle 6: Ergebnisse des Zuhöranstrengungstests mit den schwerhörigen Versuchspersonen in Experiment 2. Hier wurden die Algorithmen UN (unverarbeitet), E7, DD und DDE7 in einem Störgeräusch mit mittlerem Sprachspektrum sowie in Cafeteria-Geräusch bei 5 dB SNR benutzt.

Hearing-Impaired Subject	Step 1: Speech-Shaped Without Repetition				Step 2: Speech-Shaped With Repetition				Step 3: Cafeteria With Repetition			
	UN	E7	DD	DDE7	UN	E7	DD	DDE7	UN	E7	DD	DDE7
BD	3	4	4	5	3	4	4	3	3	2	3	2
GM	5	3	3	3	5	2	2	2	3	2	2	2
HM	5	4	3	1	5	3	4	1	4	2	5	3
KF	5	4	3	4	3	4	3	2	4	3	2	2
KR	4	5	3	3	4	5	3	3	4	4	5	3
WH	5	3	2	3	4	4	5	4	4	5	4	4
Median	5.0	4.0	3.0	3.0	4.0	4.0	3.5	2.5	4.0	2.5	3.5	2.5
MAD	0.0	0.7	0.0	0.7	1.5	0.7	0.7	0.7	0.0	0.7	2.2	0.7

3.1.2 Experiment 2

Only the hearing-impaired subjects participated in Experiment 2. Of the three different Ephraim-Malah algorithms used in Experiment 1, E7 was chosen for Experiment 2, since this algorithm performed better than EL and E30 in the adverse cafeteria noise condition.

The data suggests that *in median* the noise reduction algorithms actually *increase* listening effort rather than decreasing it (Table 6). In speech-shaped noise, algorithm E7 was judged to be the best of all the noise reduction algorithms. In median, it was one point worse than UN in Step 1, but not different to UN in Step 2. In cafeteria noise (Step 3), algorithm DD is only slightly worse than UN.

Table 7: Sentence test results of Experiment 1 with the normal-hearing subjects in drill noise. Shows the speech reception thresholds (SRT) and the slopes s of the fitted psychometric functions for the UN (unprocessed), E7, EL, and E30 algorithms. Also given are the median values and the median absolute deviations (MAD). A lower SRT corresponds to better speech intelligibility.

Tabelle 7: Ergebnisse des Satztests für die normalhörenden Versuchspersonen in Experiment 1 mit den Algorithmen UN (unverarbeitet), E7, EL und E30 unter Benutzung von Bohrmaschinenstörgeräusch. Angegeben sind die Signal-Rausch-Abstände, die zu 50 % Sprachverständlichkeit führen (SRT) und die Steigungen s der an die Testergebnisse angepassten psychometrischen Funktionen sowie Median-Werte und Medianabsolutabweichungen (MAD). Niedrigere Werte bedeuten bessere Sprachverständlichkeit.

Normal-Hearing Subject	UN		E7		EL		E30	
	SRT dB	s dB^{-1}	SRT dB	s dB^{-1}	SRT dB	s dB^{-1}	SRT dB	s dB^{-1}
AA	-21.5	0.06	-17.1	0.11	-16.7	0.10	-14.8	0.07
FJ	-22.3	0.13	-21.1	0.09	-15.2	0.06	-17.7	0.04
GI	-19.8	0.06	-21.0	0.10	-16.7	0.12	-20.2	0.04
MI	-21.6	0.07	-20.1	0.12	-19.9	0.04	-18.7	0.14
MS	-19.3	0.07	-18.3	0.10	-22.5	0.10	-19.3	0.10
RE	-21.5	0.04	-19.9	0.15	-18.2	0.08	-16.0	0.11
Median	-21.5	0.07	-20.0	0.11	-17.5	0.09	-18.2	0.09
MAD	0.7	0.01	1.6	0.01	2.2	0.03	2.3	0.05

Normal-Hearing Subject	UN		E	
	SRT dB	s dB^{-1}	SRT dB	s dB^{-1}
AA	-3.9	0.16	-2.2	0.11
FJ	-2.6	0.1	-3.4	0.15
GI	-4	0.14	-3.8	0.09
MI	-3.9	0.12	-4.1	0.23
MS	-2.7	0.15	-4.2	0.11
RE	-4.1	0.12	-2.7	0.22
Median	-3.9	0.13	-3.6	0.13
MAD	0.2	0.02	0.8	0.04

Table 8: Sentence test results of Experiment 1 with the normal-hearing subjects in cafeteria noise. E *denotes the Ephraim-Malah algorithm (E7, EL, or E30) with which each subject performed best in the drill noise. A lower SRT corresponds to better speech intelligibility.*

Tabelle 8: Ergebnisse des Satztests für die normalhörenden Versuchspersonen mit Cafeteria-Störgeräusch in Experiment 1. E *bezeichnet individuell denjenigen Ephraim-Malah-Algorithmus (E7, EL oder E30), mit dem die Versuchsperson die beste Sprachverständlichkeit im Bohrmaschinengeräusch erreichte. Niedrigere Werte bedeuten bessere Sprachverständlichkeit.*

A closer look at the data in Table 6 reveals that in Step 2, the poor median performance of the noise reduction algorithms was due to subjects GM and HM, who had strong opinions on the algorithm without noise reduction (UN). These two subjects reported that no listening effort was necessary without noise reduction, but increased effort was required with the other algorithms. Nevertheless, four out of six of the subjects still benefitted from one or the other noise reduction algorithm in speech-shaped noise (Step 2). Two subjects experienced the least listening effort with algorithm E7, one with algorithm DD, and another could not decide between E7 and DD. Two subjects judged the listening effort with DDE7 and UN to being similar, but for four subjects the listening effort increased with the combined algorithm DDE7 as compared to no noise reduction.

In cafeteria noise (Step 3), two of the six subjects experienced the least listening effort with algorithm DD, two subjects with no noise reduction (UN), one subject with E7, and one sub-

Table 9: Sentence test results of Experiment 1 with the hearing-impaired subjects in drill noise. A lower SRT corresponds to better speech intelligibility.

Tabelle 9: Ergebnisse des Satztests für die schwerhörigen Versuchspersonen mit Bohrmaschinenstörgeräusch in Experiment 1. Niedrigere Werte bedeuten bessere Sprachverständlichkeit.

Hearing-Impaired Subject	UN		E7		EL		E30	
	SRT dB	s dB^{-1}	SRT dB	s dB^{-1}	SRT dB	s dB^{-1}	SRT dB	s dB^{-1}
BD	-10.3	0.03	-10.2	0.03	-7.8	0.05	-7.9	0.10
GM	-18.0	0.10	-19.2	0.14	-17.8	0.17	-14.3	0.14
HM	-21.5	0.08	-18.2	0.17	-18.0	0.10	-16.1	0.07
KF	-23.0	0.10	-20.1	0.05	-21.7	0.07	-20.1	0.05
KR	-19.3	0.07	-19.1	0.08	-16.5	0.08	-16.0	0.11
WH	-18.6	0.05	-18.4	0.09	-13.7	0.07	-14.1	0.06
Median	-19.0	0.08	-18.8	0.09	-17.2	0.08	-15.2	0.09
MAD	2.6	0.04	0.7	0.07	3.2	0.02	1.5	0.04

Hearing-Impaired Subject	UN		E	
	SRT dB	s dB^{-1}	SRT dB	s dB^{-1}
BD	2.0	0.04	3.4	0.13
GM	-2.2	0.15	-0.1	0.11
HM	-4.6	0.12	-2.4	0.24
KF	-4.0	0.19	-2.3	0.09
KR	-3.7	0.20	-2.1	0.16
WH	-0.3	0.17	-0.3	0.19
Median	-3.0	0.16	-1.2	0.15
MAD	2.0	0.05	1.6	0.06

Table 10: Sentence test results of Experiment 1 with the hearing-impaired subjects in cafeteria noise. E *denotes the Ephraim-Malah algorithm (E7, EL, or E30) with which each subject performed best in the drill noise. A lower SRT corresponds to better speech intelligibility.*

Tabelle 10: Ergebnisse des Satztests für die schwerhörigen Versuchspersonen mit Cafeteria-Störgeräusch in Experiment 1. E *bezeichnet individuell denjenigen Ephraim-Malah-Algorithmus (E7, EL oder E30), mit dem die Versuchsperson die beste Sprachverständlichkeit im Bohrmaschinengeräusch erreichte. Niedrigere Werte bedeuten bessere Sprachverständlichkeit.*

ject could not decide between DD and UN. As regards noise reduction, three subjects had the least listening effort with DD, two with E7, and one subject was unable to decide between all three algorithms.

None of the differences between the algorithms UN, E7, DD, and DDE7 (Table 6) is statistically significant due to the low concordance among the subjects. For Step 1: $\chi^2_r = 6.288$, $df = 3$, $P = 0.098$, $W = 0.349$. For Step 2: $\chi^2_r = 6.063$, $df = 3$, $P = 0.109$, $W = 0.337$. Step 3: $\chi^2_r = 5.750$, $df = 3$, $P = 0.124$, $W = 0.319$.

3.2 Speech intelligibility

3.2.1 Experiment 1

Tables 7 to 10 show the speech reception thresholds and the slopes of the psychometric functions for the normal-hearing and hearing-impaired subjects obtained in drill noise and in cafeteria noise with the different algorithms. Also given are the median values over all subjects with the respective median absolute deviations (MAD).

In median, the noise reduction processing seems to lower speech intelligibility rather than increase it. Of the three different Ephraim-

Malah algorithms, E7 in median had the best intelligibility scores for both normal-hearing and hearing-impaired subjects.

Normal-hearing subject GI actually performed better with E7 than without noise reduction. Subject MS's SRT is 3.5 dB better with EL than without noise reduction (Table 7). The other four normal-hearing subjects, however, obtained better speech intelligibility without noise reduction.

Half of the normal-hearing subjects performed better with the noise reduction in cafeteria noise, the other half were better without. Again, subject MS benefitted most from the noise reduction with a 1.5 dB better SRT than with UN (Table 8).

Surprisingly, the normal-hearing subjects benefitted more from the noise reduction than the hearing-impaired subjects in terms of speech intelligibility. Only the hearing-impaired subject GM obtained a better speech intelligibility score in drill noise with noise reduction processing (E7) than with UN (Table 9). The hearing-impaired subjects performed worse with algorithm EL and worst with E30.

In cafeteria noise, the speech reception thresholds of the hearing-impaired subjects are on average 1.8 dB worse using the noise reduction algorithm than without noise reduction processing (Table 10). Only subject WH showed no difference between the Ephraim-Malah algorithm and UN.

A statistical analysis of the data reveals that only a few differences are actually statistically significant due to the low concordance among the subjects.

For the normal-hearing subjects in drill noise, the Friedman test results were $\chi_r^2 = 6.458$, $df = 3$, $P = 0.091$, $W = 0.359$ (no significant differences between the four algorithms; fair concordance among subjects).

Only two algorithms were tested with cafeteria noise. Hence, Wilcoxon's matched pairs signed rank test was used to test for significant differences. It yielded $Z = -0.210$, $P = 0.833$ for the normal-hearing subjects, which has no statistical significance.

For the hearing-impaired subjects in drill noise, the Friedman test yielded $\chi_r^2 = 13.068$, $df = 3$, $P = 0.004$, and this is signifciant. According to Dunn's post test, algorithms EL and E30 were significantly worse than UN. All other differences were not significant. Overall, there was substantial concordance among the subjects ($W = 0.726$). In cafeteria noise, the Wilcoxon test yielded $T = 0$, $Z = -2.023$ with an asymptotic $P = 0.043$, which is significant.[3] The hearing-impaired subjects thus obtained significantly worse SRTs with the Ephraim-Malah algorithm than without noise reduction in cafeteria noise.

[3] The exact P value is also below 0.05. Exact P values for the Wilcoxon test can, for example, be found in Table G by *Siegel* (1956).

3.2.2 Experiment 2

The results of the speech intelligibility measurements in Experiment 2 are given in Table 11. In this experiment, a speech-shaped noise was used.

Two of the six hearing-impaired subjects obtained better SRTs with algorithm E7 than with UN, three performed worse with E7, and one had almost the same SRTs with E7 and UN.

All subjects obtained better SRTs with the directional filter and dereverberation algorithm DD than with UN. The median improvement was 1.1 dB, and subject WH showed the maximum improvement (2.6 dB). The difference might seem slight, but in relation to the very steep slope of the psychometric function, it corresponds to a difference in the number of recognized test words in the order of 13 %, and in the case of subject WH, even about 26 %.

Only subject GM performed better with algorithm E7 than with DD, but her SRT differences were relatively small.

Four of the six subjects performed slightly better with the combined algorithm DDE7 than with E7. However, all subjects obtained better SRTs with algorithm DD than with DDE7. Hence, with respect to speech intelligibility, a combination of algorithms DD and E7 was not as advantageous as using DD alone.

The Friedman test results were $\chi_r^2 = 8.600$, $df = 3$ with an asymptotic $P = 0.035$ and an exact $P = 0.029$. The differences between UN and DD and between E7 and DD were found to be statistically significant. The differences between UN and algorithm E7 and between UN and DDE7 were statistically insignificant.

To summarize, the binaural algorithm DD showed a slight but significant improvement in speech reception thresholds compared to no noise reduction (UN).

4. Discussion

The listening effort test introduced here achieved its objective of assessing the differences between algorithms used on a number of normal-hearing and hearing-impaired subjects. The main characteristics of the test were rating the listening effort (according to ITU recommendations) and the step-for-step evaluation procedure that forces the subjects to really concentrate on listening effort. Anecdotally, the assessments of subject KR indicate that a repetition task as included in Step 2 really makes sense (cf. Table 5): in Step 1, subject KR reported that no effort was

Table 11: Sentence test results of Experiment 2 with the hearing-impaired subjects in speech-shaped noise using algorithms UN, E7, DD, and DDE7. A lower SRT corresponds to better speech intelligibility.

Tabelle 11: Ergebnisse des Satztests für die schwerhörigen Versuchspersonen unter Störgeräusch mit mittlerem Sprachspektrum in Experiment 2 bei Anwendung von Algorithmen UN, E7, DD und DDE7. Niedrigere Werte bedeuten bessere Sprachverständlichkeit.

Hearing-Impaired Subject	UN		E7		DD		DDE7	
	SRT dB	s dB^{-1}	SRT dB	s dB^{-1}	SRT dB	s dB^{-1}	SRT dB	s dB^{-1}
BD	-1.40	0.07	0.30	0.10	0.00	0.11	0.60	0.08
GM	-1.00	0.08	-1.90	0.12	-1.40	0.16	1.00	0.08
HM	-4.70	0.12	-3.40	0.13	-5.30	0.12	-4.30	0.09
KF	-3.00	0.14	-2.80	0.13	-3.90	0.11	-3.70	0.10
KR	-2.50	0.12	-0.70	0.16	-3.60	0.15	-0.90	0.09
WH	-0.50	0.08	-0.70	0.09	-2.10	0.13	-0.80	0.07
Median	-1.75	0.10	-1.30	0.13	-2.85	0.13	-0.85	0.09
MAD	2.59	0.03	1.56	0.02	1.85	0.02	2.45	0.01

required for any of the algorithms. In Step 2, however, the same subject was able to differentiate easily between the four algorithms. In addition to revealing some accustomization to the test situation, this might also reflect the influence of the serious listening task in Step 2, which is missing in Step 1.

Since a formal evaluation of the test procedure is still open, it has not yet been proven whether subjects are able to differentiate *reliably* between algorithms in this test. Some subjects reported that they actually experienced differences between algorithms, but that these differences did not cause them to assign different categories to the algorithms since the steps between the verbal categories were too big. Hence, it might be advisable for future studies to use an extended listening effort scale involving more categories or some subdivisions between categories. Actually, *Humes* et al. (1997), who developed a listening effort test similar to the one proposed here, required a magnitude estimate of listening effort on a 0–100 scale from their subjects in a clinical study. They used 10-sentence encyclopedia-style passages in cafeteria noise and babble backgrounds to assess the benefit of different hearing aid processing schemes. However, the test proposed by *Humes* et al. does not include a task that actually requires effort from the subjects.

Instead of using more categories in the rating procedure, paired comparisons could be applied as an alternative. As will be shown in the second paper (*Marzinzik* and *Kollmeier* 2002a), paired comparisons are very sensitive to even the slightest differences between algorithms. In general, this procedure is superior to category rating because of context and range effects and applying in a different way the scale used by the subjects in a category rating procedure (see for example *Johnson* and *Mullally* 1969). Hence, a paired comparison experiment is supposed to yield significant results even with a small number of subjects (provided differences between algorithms exist), or where category methods fail, or where many more subjects are necessary in order to show significant results (*Bech* 1987). Moreover, scale values on a difference scale level can be derived from the data obtained from paired comparisons (*Bradley* and *Terry* 1952; cf. *Marzinzik* and *Kollmeier* 2002a). However, a major disadvantage of paired comparisons is the longer time needed for measurements.

As the proposed listening effort test exposes the subject to a listening task that involves considerable effort, one expects the subject to actually *be able* to judge the listening effort after the fact. At least, one supposes that the danger of giving judgments that are greatly influenced by other perceptual dimensions such as, the »pleasantness of the sound«, »perceived artefacts« etc., will be less than in an experiment where the subject is asked to assess listening effort by merely listening (probably with hardly any effort) to a short sound sample of some seconds duration. It is believed important to let the subject really experience a situation which definitely causes or affects the phenomenon under assessment, i.e. listening effort. This is in fact the main difference between the proposed test and most of the other tests proposed in the literature so far (e.g., *Humes* et al., 1997). However, strictly speaking, there is no proof or guarantee that the proposed test actually assesses »listening effort«. In principal, it is not even *possible* to formally prove that a test actually assesses listening effort, as listening effort is assumed to be mainly a mental phenomenon (as opposed to physical fatigue, which can be assessed by muscular measurements). Some sort of proof could only be established by showing a high correlation to some »objective« measure which is again *believed* to be strongly correlated to (mental) listening effort. One example is pupillary dilation (*Hoeks* and *Levelt* 1993), another is the counting of »errors« in subsequent or parallel mental tasks. But these measures again have to be validated by subjective judgments, if a correlation to *perceived* effort is desired. Otherwise, any functional definition can only be postulated as measuring listening effort. The same problems apply to the concept of fatigue. *Muscio* (1921) concluded trenchantly that it is not possible to devise an acceptable test for fatigue because no observable criteria for fatigue exist, other than those provided by the test itself, against which the test might be validated. Whatever route is taken, one has to ask the subjects for their opinion at some point or other.

A significant improvement for hearing-impaired subjects concerning listening effort was found with the Ephraim-Malah algorithm EL compared to no noise reduction when used in drilling machine noise. In fluctuating cafeteria noise, no significant effects were found. Generally, the differences between algorithms were more pronounced for hearing-impaired subjects than for normal-hearing subjects.

Quality assessment tests (*Marzinzik* and *Kollmeier* 2002a) indicate that the Ephraim-Malah noise reduction algorithms produce more artifacts when applied in fluctuating cafeteria noise than in the stationary drilling noise. These artifacts are more prominent with algorithms EL and E30 than with E7 and seem to severely counteract possible reductions in listening effort which were expected due to signal-to-noise ratio improvements. In Experiment 2, algorithm DDE7 (the combination of E7 and DD) is characterized more by increased artifacts than by fruitful synergy effects when compared to the results of E7 and DD in isolation.

Since improvements in terms of listening effort were found with algorithm E7 in Experiment 1 but not in Experiment 2, the bad performance of E7 in the latter experiment can probably be attributed to the different noise conditions. The noise in Steps 1 and 2 was changed from drilling noise in Experiment 1 to a speech-shaped noise in Experiment 2, which more effectively masks the target speech. Moreover, the signals were deteriorated by reverberation in the second experiment.

As in Experiment 1, it was again observed in Experiment 2 that some subjects experienced no listening effort in Step 1 without noise reduction and did not benefit from the noise reduction. But in Step 2 they did and this actually required more concentration in order to fulfill the listening task. This is further evidence of the need for a really strenuous listening task if effort is to be judged correctly.

No particular correlation patterns were found between listening effort assessment, the ability to concentrate on the listening task and the age or gender of the subjects.

Future research has to address the reliability of the proposed listening effort test (i.e., determining if and how precisely subjects can reproduce their judgments in re-tests). In future experiments, it might also be worthwhile investigating the correlations between subjective preference assessments of listening effort and the length of time that a subject chose to listen to each algorithm during the experiment.

For a complete evaluation of the test itself, of course, a larger number of subjects should participate in the experiments. Then even more differences between conditions might show up as significant.

The speech intelligibility measurements confirm the picture drawn by other studies on noise suppression. Although the differences between the single-microphone noise reduction algorithms are not found to be significant, the noise reduction processing seems to decrease speech intelligibility rather than increase it. Of the three different Ephraim-Malah algorithms, E7 yields the best intelligibility scores on average for normal-hearing and hearing-impaired subjects also. It should also be noted that, on average, this algorithm does not show worse speech reception thresholds than no noise reduction (UN) with normal-hearing and hearing-impaired listeners in drill noise or in the reverberant speech-shaped noise condition. This indicates that the processing artifacts are limited in the case of the Ephraim-Malah algorithm, even though many of the enhancement systems that we are familiar with from the literature actually reduce intelligibility (*Lim* and *Oppenheim* 1979).

In cafeteria noise, however, the speech intelligibility for hearing-impaired subjects is lowered by the Ephraim-Malah noise reduction processing. This can probably be attributed to additional speech distortions which are introduced by the processing due to the fluctuating character of cafeteria babble noise. The assumption that the noise between speech pauses is stationary is strongly refuted here.

The SRTs obtained in drill noise in Experiment 1 are generally quite low compared to those obtained with speech-shaped noise in Experiment 2.This can be attributed to the fact that the drill noise has significant frequency components that are beyond even the typical speech range, and that the calculation of the signal-to-noise ratio covers the whole frequency range (0–11 kHz).

Clear improvements in terms of speech intelligibility are found with the binaural noise reduction algorithm DD in Experiment 2. These would probably not have been found if the signals had been presented dichotically (which, however, is a more realistic condition) instead of diotically. But due to the diotic presentation, the subject's own binaural processing capabilities are bypassed and the potential of the noise reduction processing itself is tested. This is supported by the fact that binaural processing capabilities vary considerably among hearing-impaired subjects and that the respective loss can not be predicted by their audiograms or other psychoacoustical parameters (*Kinkel* et al. 1991; *Kinkel* and *Kollmeier* 1992; *Holube* and *Kollmeier* 1993). Hence, the binaural system of some hearing-impaired subjects will be more effective than the directional filter and dereverberation algorithm DD. Although the amount of noise reduction can be increased, early experiments and field tests with the algorithm have shown that subjects prefer less noise reduction in favor of better overall sound quality, i.e. less artifacts (*Wittkop* 2001).

Finally, one should consider whether the lack of benefit shown with respect to speech intelligibility when single-microphone noise reduction algorithms are used could be due to a lack of acclimatization to the algorithms. *Gatehouse* (1992) found that benefits from providing a particular frequency shaping to hearing-impaired subjects did not emerge immediately, but over a period of at least 6–12 weeks. He concludes that the existence of perceptual acclimatization effects calls short-term methods of hearing aid evaluation into question. *Punch* and *Parker* (1981) point out that »*Carhart* (1946) recommended that the prospective hearing aid user be allowed to spend a substantial amount of time in individual and group listening activities prior to the recommendation of a specific instrument.«

5. Conclusions

Due to its design, which involves a strenuous listening task, the listening effort test proposed here is believed to actually assess listening effort and not merely subjective preference in terms of better sound quality. Therefore, the proposed test is recommended for the evaluations of noise reduction algorithms in general.

No increase in speech intelligibility was revealed by the Ephraim-Malah algorithms. But at least the Ephraim-Malah algorithm E7 did not make speech intelligibility worse.

Significant benefits with respect to listening effort were found for algorithm EL compared to UN (no noise reduction), although this algorithm achieved significantly worse speech reception thresholds than UN.

Acknowledgements

This work was supported by the *European Commission, Project SPACE* (Signal Processing for Auditory Communication in Noisy Environments; DE 3012) and in parts by a grant from *GN ReSound*.

References/Literatur

Baer T, Moore BCJ, Gatehouse S (1993) Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: Effects on intelligibility, quality, and response times. Journal of Rehabilitation Research 30 (1), 49–72

Bech S (1987) Planning of listening test – Choice of rating scale and test procedure. In: *Bech S, Pedersen OJ* (eds.) Perception of reproduced sound, 62–70. ISBN 87-982562-1-1

Bradley RA, Terry ME (1952) Rank analysis of incomplete block designs, I. The method of pair comparisons. Biometrika 39, 324–345

Carhart R (1946) Selection of hearing aids. Archives of Otolaryngology 44, 1–18

Davies DR, Shackleton VJ, Parasuraman R (1983) Monotony and boredom. In: *Hockey R* (ed.) Stress and Fatigue in Human Performance. John Wiley & Sons, New York, chapter 1, 1–32

Dillon H, Lovegrove R (1993) Single-microphone noise reduction systems for hearing aids: A review and an evaluation. In: *Studebaker GA, Hochberg I* (eds.) Acoustical factors affecting hearing aid performance. Allyn and Bacon, chapter 20, 353–372

Downs DW, Crum MA (1978) Processing demands during auditory learning under degraded listening conditions. J Speech Hear Res 21 (4), 702–714

Ephraim Y, Malah D (1983) Speech enhancement using optimal nonlinear spectral amplitude estimation. International Conference on Acoustics, Speech, and Signal Processing 1983. Conference Proceedings, New York, NY, USA, IEEE, 1118–1121

Ephraim Y, Malah D (1984) Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator. IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP-32 (6), 1109–1121

Ephraim Y, Malah D (1985) Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP-33 (2), 443–445

Fabry DA, Schum DJ (1994) The role of subjective measurement techniques in hearing aid fittings. In: *Valente M* (ed.) Strategies for selecting and verifying hearing aid fittings. Thieme Medical Publishers, New York, 136–155

Gatehouse S (1992) The time course and magnitude of perceptual acclimatization to frequency responses: Evidence from monaural fitting of hearing aids. J Acoust Soc Am 92 (3), 1258–1268

Gatehouse S (1994) Components and determinants of hearing aid benefit. Ear Hear 15, 30–49

Hecker MHL, Stevens KN, Williams CE (1966) Measurements of reaction time in intelligibility tests. J Acoust Soc Am 39 (6), 1188–1189

Hoeks B, Levelt WJM (1993) Pupillary dilation as a measure of attention: A quantitative system analysis. Behavior Research Methods, Instruments, & Computers 25 (1), 16–26

Holding DH (1983) Fatigue. In: *Hockey R* (ed.) Stress and fatigue in human performance. John Wiley & Sons, New York, chapter 6, 145–167

Holube I, Kollmeier B (1993) A perception model to predict speech intelligibility in impaired listeners using psychoacoustical parameters. In: *Schick A* (ed.) Contributions to Psychological Acoustics, 6th Oldenburg Symposium, Oldenburg. BIS Oldenburg, 557–566

Hornik K (2000) The R FAQ. http://www.ci.tuwien.ac.at/~hornikR/

Humes LE, Christensen LA, Bess FH, Hedley-Williams A (1997) A comparison of the benefit provided by well-fit linear hearing aids and instruments with automatic reductions of low-frequency gain. J Speech Hear Res 40, 666–685

Hygge S, Rönnberg J, Larsby B, Arlinger S (1992) Normal-hearing and hearing-impaired subjects' ability to just follow conversation in competing speech, reversed speech, and noise backgrounds. J Speech Hear Res 35, 208–215

Ihaka R, Gentleman R (1996) R: A language for data analysis and graphics. Journal of Computational and Graphical Statistics 5 (3), 299–314

ITU (1996) ITU-T Recommendation P.800: Methods for subjective determination of transmission quality. International Telecommunication Union.

Ivarsson US, Arlinger SD (1993) Speech recognition in noise before and after a work-day's noise exposure. Scand Audiol 23, 159–163

Johnson DM, Mullally CR (1969) Correlation-and-regression model for category judgments. Psychological Review 76 (2), 205–215

Jones DM (1983) Noise. In: *Hockey R* (ed.) Stress and fatigue in human performance. John Wiley & Sons, New York, chapter 3, 61–95

Kinkel M, Kollmeier B (1992) Binaurales Hören bei Normal- und Schwerhörigen II: Analyse der Ergebnisse. Audiologische Akustik 31, 22–33

Kinkel M, Kollmeier B, Holube I (1991) Binaurales Hören bei Normalund Schwerhörigen I: Methoden und Ergebnisse. Audiologische Akustik 30, 192–201

Kollmeier B, Peissig J, Hohmann V (1993) Binaural noise-reduction hearing aid scheme with real-time processing in the frequency domain. Scand Audiol Suppl 38, 28–38

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 (4), 2412–2421

Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33, 159–174

Levitt H, Bakke M, Kates J, Neuman A, Schwander T, Weiss M (1993) Signal processing for hearing impairment. Scand Audiol Suppl 38, 7–19

Lim JS, Oppenheim AV (1979) Enhancement and bandwidth compression of noisy speech. Proceedings of the IEEE 67 (12), 1586–1604

Lybarger S (1944) Method of fitting hearing aids. U.S. Patent Application S.N. 543, 278

Lybarger S (1978) Selective amplification – a review and evaluation. J Am Audio Soc 3 (6), 258–266

Marascuilo LA, McSweeney M (1977) Nonparametric and distribution-free methods for the social sciences. Brooks/Cole, Monterey, California

Marzinzik M, Kollmeier B (2001) A review of the Ephraim-Malah noise reduction algorithms. Z Audiol 40 (1), 4–15

Marzinzik M, Kollmeier B (2002a) Noise reduction schemes for digital hearing aids: II. Subjective quality assessment based on paired comparisons. Z Audiol submitted

Marzinzik M, Kollmeier B (2002b) Speech pause detection for noise spectrum estimation by tracking power envelope dynamics. IEEE Transactions on Speech and Audio Processing 10 (2), 109–118

Motulsky HJ (1995) Intuitive Biostatistics. Oxford University Press

Muscio B (1921) Is a fatigue test possible? British Journal of Psychology 12, 31–46. Cited in: *Holding* (1983)

Peissig J (1993) Binaurale Hörgerätestrategien in Störschallsituationen. VDI-Verlag, Düsseldorf

Punch J, Parker C (1981) Pairwise listener preferences in hearing aid evaluation. J Speech Hear Res 24, 366–374

Roethlisberger FJ, Dickson WJ (1939) Management, and the Worker. Harvard University Press, Cambridge, MA. Cited in: *Jones* (1983)

Siegel S (1956) Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill, New York

Studebaker GA, Hochberg I (1993) Acoustical factors affecting hearing aid performance. Allyn and Bacon

Wagener K, Kühnel V, Brand T, Kollmeier B (1998) Entwicklung und Evaluation eines Sprachverständlichkeitstests für die deutsche Sprache. Fortschritte der Akustik – DAGA 98, Oldenburg, DEGA, 322–323. ISBN 3-9804568-3-8

Wagener K, Kühnel V, Kollmeier B (1999) Entwicklung und Evaluation eines Satztests für die deutsche Sprache. Z Audiol 38 (1, 2, 3)

Weiss M, Neuman AC (1993) Noise reduction in hearing aids. In: *Studebaker GA, Hochberg I* (eds.) Acoustical factors affecting hearing aid performance. Allyn and Bacon, chapter 19, 337–352

Wittkop T (2001) Two-channel noise reduction algorithms motivated by models of binaural interaction. Carl von Ossietzky Universität Oldenburg. Doctoral dissertation

Wittkop T, Hohmann V, Kollmeier B (1999) Noise reduction strategies employing interaural parameters. ACUSTICA Acta Acustica 85, 285

Wright SP (1992) Adjusted p-values for simultaneous inference. Biometrics 48, 1005–1013

Wyatt S (1950) An autobiography. Occupational Psychology 24, 65–74. Cited in: *Davies* et al. (1983), p. 5

Zwicker E, Terhardt E (1980) Analytical expressions for critical-band rate and critical bandwidth as a function of frequency. J Acoust Soc Am 68 (5), 1523–1525

ORIGINALARBEIT

Evaluation of different 3-channel dynamic compression schemes in a field test with a wearable DSP prototype hearing aid

*Jens E. Appell *+***, Volker Hohmann *, Birgitta Gabriel **, Birger Kollmeier **
* Medizinische Physik, Universität Oldenburg, D-26111 Oldenburg
** Hörzentrum Oldenburg, D-26111 Oldenburg
*** Current Address: OFFIS – Embedded Systems, Escherweg 2, D-26121 Oldenburg

Abstract *Three different 3-channel dynamic compression schemes (automatic volume control, syllabic compression and compression limiting), previously tested in the laboratory by Appell et al. (1995), were investigated in everyday life by five hearing-impaired subjects using a binaural prototype digital hearing aid. A battery of tests was performed that included categorical loudness scaling with narrow- and broadband stimuli, speech intelligibility in noise using the Göttingen Sentence Test and quality assessment by paired comparison tests as well as a questionnaire and an informal interview. Due to the carefully selected control conditions (i. e., unaided situation at roughly the same perceived loudness as in the aided situations, algorithms fitted with the same fitting rationale, same frequency response for all algorithms at a medium input level with a speech-shaped spectrum) the differences across algorithms were only very slight. Hence, no overall »winner algorithm« could be derived from the data. However, the results showed that subjects with a low residual dynamic range and high speech reception thresholds (SRT) showed best performance in quality and speech intelligibility with dynamic compression whereas no clear-cut preference is found in the other subjects. The current study suggests that slow-acting compression with a high compression ratio (i.e., automatic volume control) should be used for low input levels to provide audibility at this input level range, whereas syllabic compression (small compression ratio) or even linear amplification seems to be beneficial at medium to high input levels. In any case, compression limiting should be provided to prevent high-level signal peaks.*

Key words: *hearing aids*
dynamic compression
AGC (automatic gain control)

Corresponding author: Dr. rer. nat. Dipl.-Phys. Jens-E. Appell
OFFIS – R&D Division Embedded Systems
Escherweg 2
D-26121 Oldenburg
Phone +49 441 9722 286
Fax +49 441 9722 282
E-mail: jens-e.appell@offis.de

Feldtest verschiedener Dreikanal-Dynamikkompressionsalgorithmen mit einem tragbaren digitalen Hörgeräteprototypen

*Jens E. Appell *+***, Volker Hohmann *, Birgitta Gabriel **, Birger Kollmeier **
* Medizinische Physik, Universität Oldenburg, D-26111 Oldenburg
** Hörzentrum Oldenburg, D-26111 Oldenburg
*** Current Address: OFFIS – Embedded Systems, Escherweg 2, D-26121 Oldenburg

Zusammenfassung *Drei unterschiedliche Strategien zur Dynamikkompression (Automatic-Volume-Control: AV-Algorithmus, Silbenkompression: SC-Algorithmus und lineare Verarbeitung mit compression limiting: CL-Algorithmus), die bereits in Laborexperimenten getestet wurden (Appell et al. 1995), wurden mit Hilfe eines binauralen digitalen dreikanaligen Hörgeräteprototypen mit fünf Schwerhörenden im Alltag getestet. Zusätzlich wurde die Lautheitswahrnehmung mit der kategorialen Lautheitsskalierung für schmalbandige und breitbandige Signale bestimmt und das Sprachverstehen mit Hilfe des Göttinger Satztests ermittelt. Die Verarbeitungsqualität der Algorithmen wurde im Paarvergleich getestet und mit Hilfe eines Fragebogens und eines informellen Interviews erfragt. Aufgrund der unvoreingenommenen Wahl der Versuchsbedingungen (z. B.: Darbietung bei gleicher empfundener Lautheit in der Situation versorgt und unversorgt, Verwendung der gleichen Anpassstrategie für alle Algorithmen, gleicher Frequenzgang für ein sprachähnliches Eingangsspektrum bei mittleren Pegeln) konnten nur geringe Unterschiede zwischen den Algorithmen festgestellt werden, so dass kein »Gewinner-Algorithmus« ermittelt werden konnte. Dennoch zeigen die Ergebnisse, dass Patienten mit einer geringen Restdynamik und hohen Sprachverständlichkeitsschwellen (SRT) eine Dynamikkompression als qualitativ besser beurteilen und mit Dynamikkompression eine bessere Sprachverständlichkeit erreichen als ohne. Bei den Patienten mit einer höheren Restdynamik konnte keine klare Preferenz für einen der Algorithmen festgestellt werden. Die Ergebnisse legen nahe, dass bei niedrigen Eingangspegeln eine langsam arbeitende Kompression mit großen Kompressionsfaktoren (automatic volume control) eingesetzt werden sollte, um die Hörbarkeit solcher Schalle sicher zu stellen. Dagegen erscheint eine Silbenkompression oder eine lineare Verarbeitung bei mittleren bis hohen Pegeln vorteilhaft. Ein Schutz vor zu lauten Schallen (Lautstärkespitzen) ist in jedem Fall erforderlich und kann beispielsweise durch Compression-Limiting erreicht werden.*

Schlüsselwörter: Hörgeräte
Dynamikkompression
AGC (automatic gain control)

1. Introduction

One of the most common principles in modern digital hearing-aids is dynamic range compression. It is introduced in order to compensate for the so-called recruitment phenomenon (*Steinberg* and *Gardner* 1937), i.e. deteriorated loudness perception in sensorineural hearing-impaired patients. While the classic solution is a broadband automatic volume control (AVC or gain adjustment of the hearing-aid across the full frequency range according to the input level averaged over a certain time window), more advanced multiband dynamic compression systems have been suggested and evaluated the past years (i.e., *Lippmann* et al. 1981; *NáběIek* 1983; *Walker* and *Byrne* 1984; *Genaro* et al. 1986; *White* 1986; *Bustamente* and *Braida* 1987; *Plomp* 1988; *Fröhlich* 1993; *Hohmann* and *Kollmeier* 1995a; *Marzinzik* et al. 1997; *Tejero-Calado* et al. 1998; *Moore* et al. 1999; *Festen* 1999; *Stone* et al. 1999). A comprehensive overview is given by the *Working Group on Communication Aids for the Hearing-Impaired* (1991), *Kollmeier* (1997a) and *Verschuure* and *Dreschler* (1996). The aim of these compression algorithms is to restore a maximum number of the impaired auditory functions found in hearing-impaired patients in the best possible way. Given the limited signal processing capabilities available in hearing-aids and our limited knowledge about »effective« signal processing of the normal and impaired auditory system, this restoration can only be incomplete. Therefore, an important partial goal is to at least compensate for the altered loudness impression in hearing-impaired listeners and/or to provide the optimum presentation level of the input signal. This should help the hearing-impaired listener to process speech in an optimum way in order to optimize speech intelligibility.

In any case, linear frequency shaping is required, typically with a fine spectral resolution, in order to equalize the frequency response independently from the input level. In addition, a nonlinear compression component is required in order to compress the large dynamic range of input signals to the limited dynamic range of the impaired ear. Typically, the latter operation can be performed at a lower frequency resolution than linear frequency shaping. Several multi-band dynamic compression schemes have been suggested that usually perform dynamic compression in several frequency bands independently. In general, the sound quality (and, in most cases, also the performance in terms of restoring speech intelligibility in quiet and in noise) deteriorates with an increasing number of frequency channels and with decreasing time constants (*Festen* 1999; *Genaro* et al. 1986; *Hansen* 2000; *Plomp* 1988; *NáběIek* 1983; *Neuman* et al. 1995). However, evaluation studies on dynamic compression hearing-aids typically have the following shortcomings:

- The linear frequency shaping (which should be applied at a medium input level and provide the frequency shaping normally introduced by fitting rules for linear hearing-aids such as NAL (*Byrne* 1986)) is typically not separated from the nonlinear component of the level correction (i.e., compression in a few bands). Hence, when different compression schemes are being compared with each other, it is necessary to ensure that all the algorithms provide the same frequency shaping for a certain input signal (such as speech-simulating noise) at a certain input level (e.g., medium conversation level).

- The evaluation is usually only based on laboratory studies with a limited set of acoustical situations and a very limited set of input levels. In daily life, however, dynamic compression algorithms have to perform under a variety of acoustical conditions and at different presentation levels. Hence the need for a more comprehensive comparison of dynamic compression systems that encompasses a wide range of conditions and levels as well as a field test.

- Most studies have concentrated on providing comparatively few items for hearing-impaired listenersto assess (e.g., speech intelligibility in quiet and in noise). However, other important items that characterize hearing-aid performance in real-life situations should be considered (e.g., loudness compensation for a variety of input levels and bandwidths, subjective assessment of the hearing-aid and overall quality rating).

Hence, the current study tries to provide a valid comparison of different compression schemes for multi-channel dynamic compression hearing-aids (syllabic compression, automatic volume control, and compression limiting), concentrating on a variety of different evaluation criteria and using a wearable prototype signal-processing aid in a field test. The aim of the study is to find out – under controlled experimental conditions, both in the laboratory and in the field – if there are any consistent differences across compression rationales and fitting rules.

2. Processing Schemes

Within this study, several different 3-channel dynamic compression schemes were investigated in everyday situations. They had all previously been tested in the laboratory by *Appell* et al. (1995). The main parameters of the processing schemes were two cutoff frequencies separating the three frequency channels, the attack and release-time constants of the input level estimators and the input-output characteristics (I/O characteristic) of the dynamic compression. A schematic overview of the 3-channel master hearing-aid algorithm employed in this study is given in Figure 1.

The algorithms were implemented on a wearable digital hearing-aid device, the so-called *DASi-2* (Digital Auditory Signal-processor, version 2) developed by *Raß* and *Steeger* (2000). The *DASi-2* consists of two Siemens ITE devices (Cosmea M) connected binaurally to a pocket-size signal processor device. The

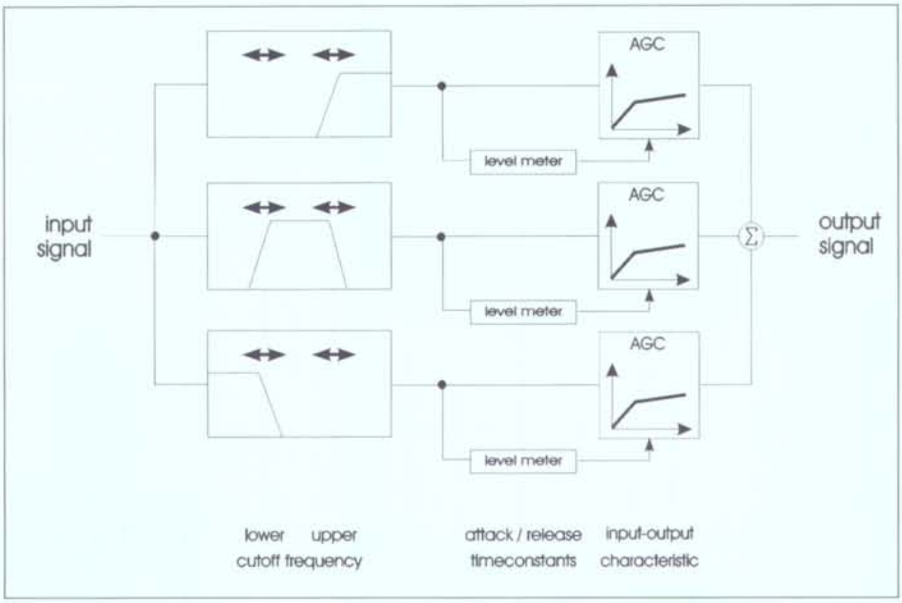


Fig. 1: Schematic overview of the 3-channel dynamic compression master hearing-aid.

Abb. 1: Blockschaltbild des dreikanaligen Dynamikkompressionsalgorithmus. Die unterschiedlichen Verarbeitungsstrategien wurden durch die Wahl entsprechender Parameter realisiert (Master Hearing-Aid Konzept).

signal processing framework implemented on the DASi provides signal processing in the frequency domain using an Overlap-Add processing scheme (18.9 kHz sampling frequency, 180 sample Hanning windowing with 50 % overlap, zero padding by 76 sample resulting in 256 point FFT).

The algorithms were implemented as follows: the low-pass, band-pass and high-pass signals are obtained by summing up the intensities $I_n(f)$ of the FFT bins $f_{start}(c)$ to $f_{end}(c)$ belonging to the respective channel c:

$$I_n(c) = \sum_{f_{start}(c)}^{f_{end}(c)} I_n(f) \tag{1}$$

where n is used as the sample index. The two cutoff frequencies f_{LowCut} and $f_{HighCut}$ between the three channels are adjustable by selecting the corresponding FFT bins:

$$\begin{aligned} f_{LowCut} &= f_{end}(c=1) = f_{start}(c=2)-1 \\ f_{HighCut} &= f_{end}(c=2) = f_{start}(c=3)-1 \end{aligned} \tag{2}$$

The estimated band signal levels are then calculated from the respective summed intensities by applying a temporal first-order recursive averaging filter and subsequently transforming it into the dB scale. The averaging filter allows for the definition of different attack and release time constants for each AGC channel, i.e., different adaptation times for rising and falling input levels as follows: in the first step, a peak hold with decay applies the release time constant $\tau_{rel}(c)$ when the input level decreases:

$$I_n(c) = \begin{cases} I_n(c) & \text{for } I_n(c) \geq \tau_{rel}(c)\cdot L_{n-1}(c) \\ \tau_{rel}(c)\cdot L_{n-1}(c) & \text{for } I_n(c) < \tau_{rel}(c)\cdot L_{n-1}(c) \end{cases} \tag{3}$$

Then gain $G_n(c)$ is calculated from an I/O characteristic (see below) using $L_n(c)$ as the input parameter. Subsequently, gain $Gn(c)$ is smoothed by a first order IIR filter[1], applying the attack time constant $\tau_{att}(c)$:

$$\overline{G_n(c)} = \tau_{att}(c)\cdot\overline{G_{n-1}(c)} + (1-\tau_{att}(c))\cdot G_n(c) \tag{4}$$

Within each frequency channel, an I/O characteristics is defined that prescribes the desired output level as a function of the estimated signal level on a log-log scale. The current gain in each band is then calculated from the respective input level using the I/O characteristics and applied to the band signals. The output signal is formed by summing up the modified band signals. In this way, the frequency channels are compressed independently of each other.

The I/O characteristic uses 3 piece-wise linear sections, i.e., 2 kneepoints. It is implemented as a table lookup with linear interpolation between the table entries. Figure 2 shows examples of the compression characteristics for different compression schemes.

Based on the 3-channel master compression algorithm described above, 4 different compression schemes (i.e., settings of its parameters) were defined, representing fundamentally different approaches for dynamic range reduction. Note that all algorithms were adjusted so as to provide the same (average) amplification for a speech spectrum shaped signal at 65 dB SPL, (i.e., the level that corresponds to the average MCL for normal listeners for speech signals). The different schemes are described as follows:

[1] IIR: infinite impulse response

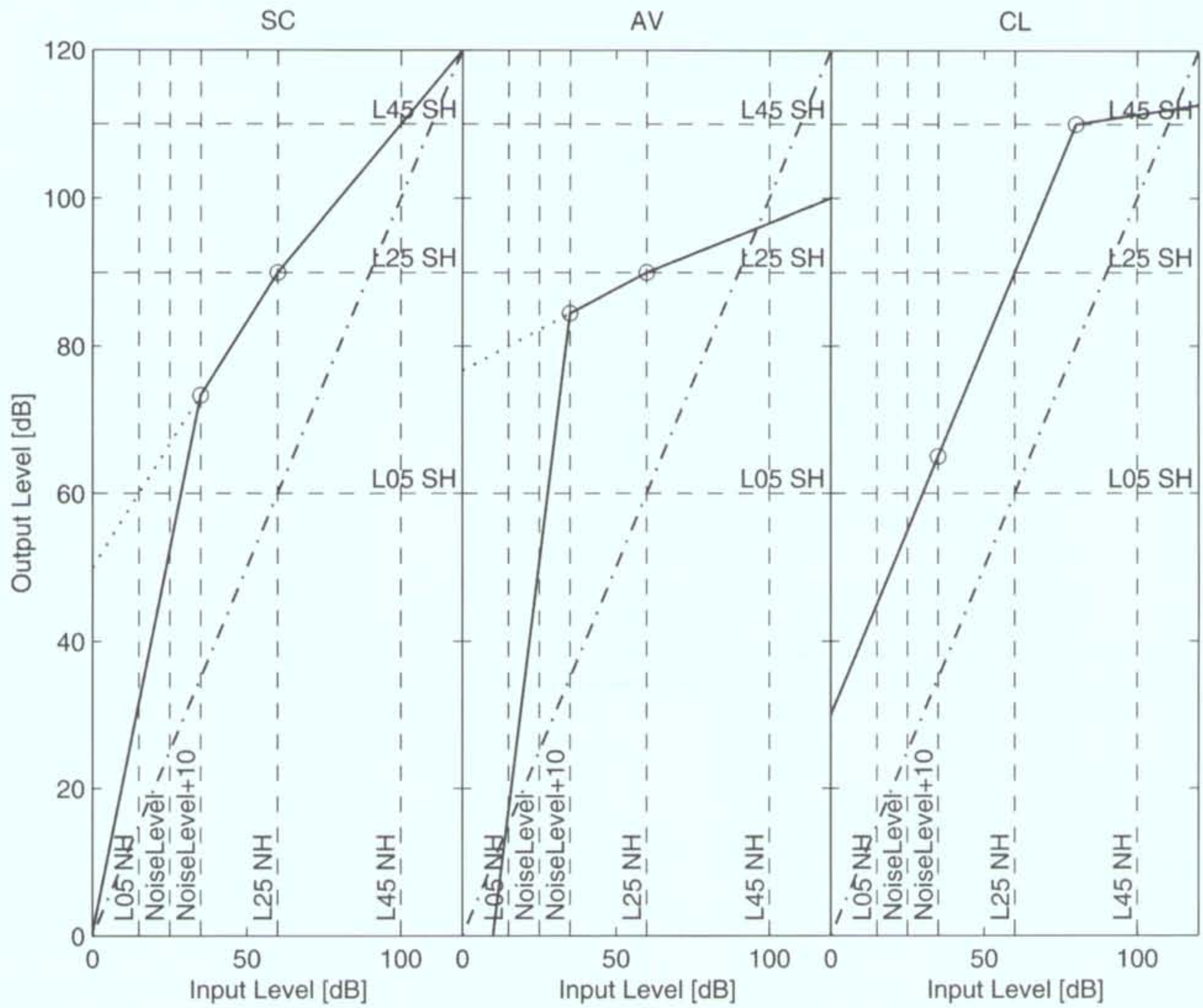


Fig. 2: Example of the I/O characteristics (prescribed output level as a function of input level on a dB-scale) for one frequency channel of the dynamic compression algorithms SC, AV and CL. The solid lines show the compressive and expansive parts of the I/O characteristics, kneepoints are shown by circles (○). The dotted lines indicate the continuation of the I/O characteristic when no expansion was implemented. The dash-dotted line shows a linear system without amplification. The vertical and horizontal lines show the levels corresponding to »very soft«, »medium« and »very loud« (L05, L25 and L45, respectively) for normal-hearing subjects (NH, vertical lines) and hearing-impaired listeners (SH, horizontal lines), as well as the definition of the lower kneepoint at 10 dB above the noise oor of the device.

Abb. 2: Ein Beispiel für die I/O-Kennlinien (Ausgangspegel als Funktion des Eingangspegels in dB) eines Frequenzkanals für die Algorithmen SC, AV und CL (von links nach rechts). Die durchgezogenen Kurven zeigen die kompressiven und expansiven Abschnitte dieser Kennlinien, Kreise (○) kennzeichnen die Knickpunkte. Die punktierten Kurven zeigen die Fortsetzung der Kennlinie ohne Expansionsteil. Die strichpunktierten Kurven zeigen ein lineares System ohne Verstärkung. Die vertikalen und horizontalen Kurven zeigen die Pegel, die der kategorialen Lautheitswahrnehmung »sehr leise«, »mittellaut« und »sehr laut« (L05, L25 und L45) bei Normalhörenden (NH, vertikale Linien) und eines Schwerhörenden (SH, horizontale Linien) entsprechen. Die mit »NoiseLevel« und »NoiseLevel+10« gekennzeichneten vertikalen Linien entsprechen dem Pegel des Grundrauschens des Hörgerätes und einem Pegel 10 dB darüber.

Linear amplification (*LIN*)

The linear gain applied by the linear amplification scheme denoted as *LIN* is based on a mapping of the unaided *MCL* (most comfortable level, category »medium« in the loudness scaling, see 5.2) to the average MCL of normal-hearing listeners. This linear frequency shaping was performed in 20 non-overlapping frequency bands with a bandwidth of 1 critical band (1 Bark). Since this scheme does not include dynamic compression, the patient is not protected from very loud sounds at high input levels. Therefore, this scheme was only used as a reference in the laboratory experiments.

Compression limiting (*CL*)

This algorithm is based on algorithm *LIN* but additionally provides compression limiting at high input levels and can therefore protect the patient from overly loud signals (see Figure 2, right panel). Compression limiting was realized using a compression ratio (i.e., the inverse of the slope of the I/O characteristics) of 15, attack time constants of 1 msec and release time constants of 50, 7 and 3 msec in the low-pass, band-pass and high-pass channels, respectively, and a compression kneepoint corresponding to the output level that matched the judgement »very loud« in the loudness-scaling experiment. This algorithm was one program selected for the tests in everyday life.

Syllabic compression (*SC*)

The syllabic compression used the same linear frequency shaping as *LIN* and *CL*, i.e., the same mapping of the unaided MCL to normal-hearing MCL. In addition, a compressive I/O characteristic was implemented which matches loudness at low (category »very soft«) and high (category »very loud«) levels. Since the high gain produced by the dynamic compression at low input levels would make the noise floor of the audio hardware audible to the patient, an expansive part was implemented in the I/O characteristic at very low input levels (see Figure 2, left panel). The kneepoint for the expansive characteristics was set 10 dB higher than the noise level of the audio hardware in each channel. In addition, the compression ratio was set so that input levels equal to the noise level of the audio hardware had the same amplification as the overall gain at normal-hearing MCL (L25 NH in Figure 2). The frequency-channel-dependent time constants used were short enough to follow the frequency of syllables (envelope compression) and long enough not to distort the signal's waveform. Therefore, the same attack and release time constants of 50, 7 and 3 msec in the low-pass, band-pass and high-pass channel, respectively, were chosen as in the compression limiting case. This algorithm was one program selected for the tests in everyday life.

Automatic volume control (*AV*)

The aim of this specific compression algorithm was to compress the input signal »effectively« by an amount similar to that of algorithm *SC*, but not to compress the temporal structure of the input signals in order to preserve the temporal contrasts. Since the »effective« compression of speech signals is a function of the modulation transfer function for the compressive system and the modulation spectrum of speech (c.f., *Hohmann* and *Kollmeier* 1995a; *Plomp* 1988; *Villchur* 1989; *Verschuure* and *Dreschler* 1996), larger time constants have to be counteracted by a higher compression ratio in order to still yield the same »effective« compression. Hence, the individual compression ratios obtained for each subject by algorithm *SC* were increased by a factor of 3 and longer attack and release time constants of 200 msec were used in all channels. Whereas the static compression characteristic (cf. Fig. 2) indicates that the maximum output level for stationary signals is limited, this does not hold for fluctuating speech signals due to the relatively long time constants. The other parameters chosen were the same as for algorithm *SC*. This algorithm was one program selected for the tests in everyday life.

3. Subjects

Five sensorineurally hearing-impaired patients with mild to moderate hearing loss were selected for the tests. All patients are experienced hearing-aid users and were motivated as well as skilled enough to handle the prototype DSP hearing-aid. They received a nominal fee for their participation in the study.

At the beginning of the field test, each subject participated in a complete routine audiological examination, including a pure tone audiogram determination of bone conduction hearing loss and Uncomfortable Loudness Level (*UCL*) as well as impedance audiometry. The pure tone audiogram (*PTA*), the bone conduction hearing loss and the UCL were measured using a KIND DA930 audiometer (Interacoustics AC30). The air conduction threshold was measured for each ear at frequencies of 0.125, 0.25, 0.5, 1, 2, 3, 4, 6 and 8 kHz. The bone conduction threshold and UCL were determined at 0.5, 1, 2, 3 and 4 kHz. Table 1 shows the data of the five subjects participating. The difference between air and bone conduction was less than 15 dB at all frequencies tested (0.5, 1, 2, 3 and 4 kHz).

Table 1: Pure tone audiogram (air conduction thresholds), bone conduction hearing loss and uncomfortable loudness level in dB HL for all subjects. The mean value and standard deviation are given in the last two rows for all frequencies, respectively. Data denoted by »nm« could not be gathered due to the limited output level of the audiometers. The symbol »-« indicates situations that were not measured.

Tab. 1: Luftleitungs-, Knochenleitungs- und Unbehaglichkeitsschwellen der Probanden in dB HL. Mittelwerte und Standardabweichung sind in den unteren zwei Zeilen gegeben. Die mit »nm« angegebenen Einträge kennzeichnen die Meßwerte, die aufgrund der Limitierung der Audiometerverstärkung nicht ermittelt werden konnten. Situationen, die nicht gemessen wurden, sind mit »-« gekennzeichnet.

Sub	Ear	250 Hz	500 Hz	1000 Hz	2000 Hz	3000 Hz	4000 Hz	6000 Hz
BD	right	45/ -/ -	50/50/105	50/50/110	30/30/110	40/40/110	55/45/110	65/ -/ -
BD	left	55/ -/ -	55/55/110	55/50/105	45/45/105	70/55/105	70/60/110	80/ -/ -
EJ	right	45/ -/ -	55/55/nm	55/60/120	55/60/115	50/50/nm	50/50/nm	50/ -/ -
EJ	left	55/ -/ -	60/50/nm	60/60/120	45/60/110	50/50/110	45/55/nm	45/ -/ -
GH	right	75/ -/ -	70/55/105	70/55/105	70/65/110	70/70/115	75/70/115	90/ -/ -
GH	left	50/ -/ -	40/35/ 95	40/35/100	35/25/ 95	35/30/ 95	40/30/100	45/ -/ -
HM	right	25/ -/ -	35/40/ 90	55/50/ 95	55/55/ 90	45/45/ 85	60/60/ 90	65/ -/ -
HM	left	20/ -/ -	35/30/ 90	50/45/ 85	50/50/ 85	55/50/ 90	60/55/ 90	70/ -/ -
MW	right	50/ -/ -	55/45/ 95	70/60/100	80/65/100	70/55/100	70/55/100	70/ -/ -
MW	left	55/ -/ -	55/45/ 95	65/50/100	75/70/100	75/65/100	80/65/105	80/ -/ -
Mean		48/ -/ -	51/46/ 98	57/52/104	54/53/102	56/51/101	61/55/103	66/ -/ -
STD		16/ -/ -	11/ 9 / 8	9 / 8 / 11	17/15/ 10	14/11/ 10	13/11/ 9	15/ -/ -

The mean hearing loss across all frequencies in table 1 was about 55 dB HL. It increased slightly with frequency and did not vary much among subjects. The UCL of all subjects was recorded at about 100 dB HL for all frequencies between 0.5 and 4 kHz, which corresponds to the UCL of normal-hearing listeners. Hence, the hearing-impaired listeners showed recruitment with a residual dynamic range of about 40 to 70 dB.

All five subjects participated in the laboratory tests, but only four subjects (i.e., subject EJ excluded) participated in the field test.

4. Fitting of the Compression Schemes to the Individual Hearing Loss

In the first step of the fitting procedure, the linear gain is adjusted to obtain a mapping of the unaided MCL to the average MCL of normal-hearing listeners. The data required for this prescriptive step of the fitting was measured by monaural loudness scaling using narrow-band noise signals (see section 5.2). Figure 3 illustrates the components of the wearable hearing-aid which determine the amplification of the device. The microphone amplifier in the *Cosmea M* ITE hearing-aid was set to its maximum value (vAmp = 17 dB) in order to obtain a maximal signal level on the rather long wires leading to the digital device. The analog attenuator $V_{A/D}$ in front of the *ADC* (analog to digital converter) was adjusted to yield a digital input level of -12 dB relative to the overload level of the ADC for a 90 dB SPL speech simulating noise input signal. This was done in order to shift the effective dynamic range of the ADC (approximately 90 dB) to signal levels common to everyday situations. A first rough fitting to an individual hearing loss was achieved by setting the attenuation of the DAC ($V_{D/A}$) to a value which ensured the desired amplification in the speech-relevant range of 500 to 4000 Hz. Fine tuning was done by setting further frequency-dependent attenuation in the digital domain for each frequency channel (critical band) or – if required in order to achieve the prescribed gain – on the basis of FFT Bins. Afterwards, a loudness scaling of narrow-band noises was performed in the aided condition (center frequencies 0.5, 1

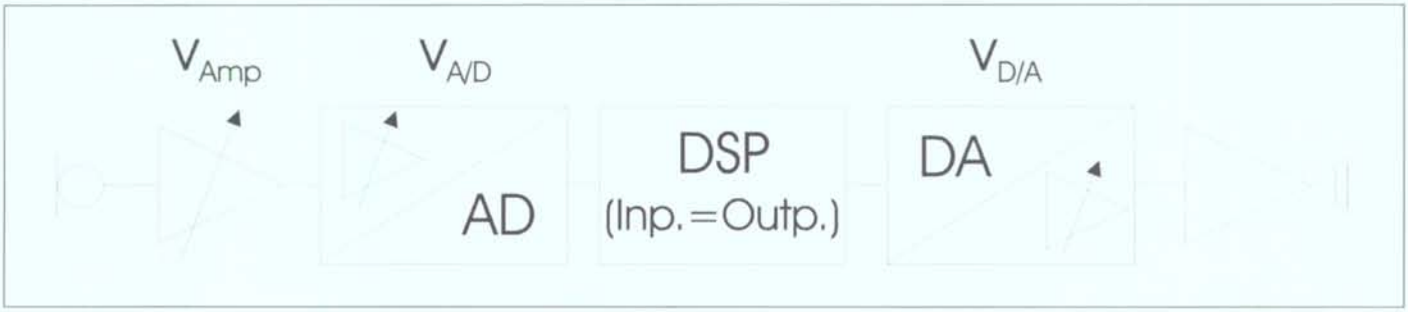


Fig. 3: Block diagram of the signal ow in the DASi-2 (Raß and Steeger 2000).

Abb. 3: Blockdiagramm des Signalweges des tragbaren Digitalen Hörgerätes DASi-2 (Raß and Steeger 2000).

and 3 kHz) in order to verify the gain setting. Where the MCL in the aided condition deviated more than 10 dB from the average MCL of normal-hearing listeners, the amplification in the corresponding frequency channel was readjusted and the loudness scaling repeated[2]. Finally, a loudness scaling with speech-shaped noise was carried out to adjust the overall level for broadband signals. It was set so that MCL in the aided condition would be achieved at 65 dB SPL. Because broadband signals in general are judged to be louder than narrow band signals at the same physical level (due to spectral loudness summation), this correction led to a slight attenuation of a few decibels towards lower levels and was achieved by attenuating the output of the speech processor digitally.

Within the field test the volume control switch on the digital device was programmed to let the user increase the prescribed amplification by up to 6 dB and reduce it down to -18 dB. In all laboratory experiments the volume control switch was set to 0 dB.

In order to fit the dynamic compression algorithms described in section 2 to the individual hearing loss, the I/O characteristics in the three frequency channels and the cutoff frequencies between these channels had to be adjusted. Since all subjects showed a relatively at audiogram, the cutoff frequencies could be set to the same values of 703 Hz and 1828 Hz for all subjects. These cutoff frequencies provide a nearly equal distribution of the input (speech) energy to each of the three channels and were suggested by *Kießling* and *Steffens* (1991) for other three-channel dynamic compression systems.

The I/O characteristics of the dynamic compression algorithms under study were determined by the loudness scaling data. However, because of the fixed point representation of the signal within the DASi's DSP, the maximum gain in the gain table defining the I/O characteristic could not exceed 24 dB. Therefore, the dynamic compression algorithms – especially algorithm *AV* – could not be fitted exactly as prescribed. In addition, it turned out that the *AV* algorithm had a tendency to produce feedback at low input levels due to its high gain. In both cases, the maximum gain of the I/O characteristics had to be reduced. Figure 4 shows an example where extreme limitations had to be made. It should be pointed out that these restrictions did not have to be made in the laboratory study by *Appell* et al. (1995), where the same algorithms were tested on a stationary signal processing system[3].

Throughout the field test, four programs were stored on the DASi. Thus, the subject could easily switch between four hearing-aid algorithms. Table 2 lists the algorithms assigned to the four programs. In addition to the 3 processing schemes described here, the field test was carried out with one additional algorithm providing a combination of dynamic compression with a noise suppression algorithm. The results of this algorithm are reported elsewhere (*Appell* et al. 1999).

5. Assessment Methods

A battery of audiological tests was used in order to measure speech intelligibility, loudness perception and the system's sound quality as well as the acclimatization effects.

[2] In general no readjustments had to be made here.

[3] In the laboratory study by *Appell* et al. (1995), the input levels of the stimuli were well-known, so the dynamic range of the system could be adjusted accordingly. In addition, all stimuli were presented via headphones, thus preventing feedback in the system.

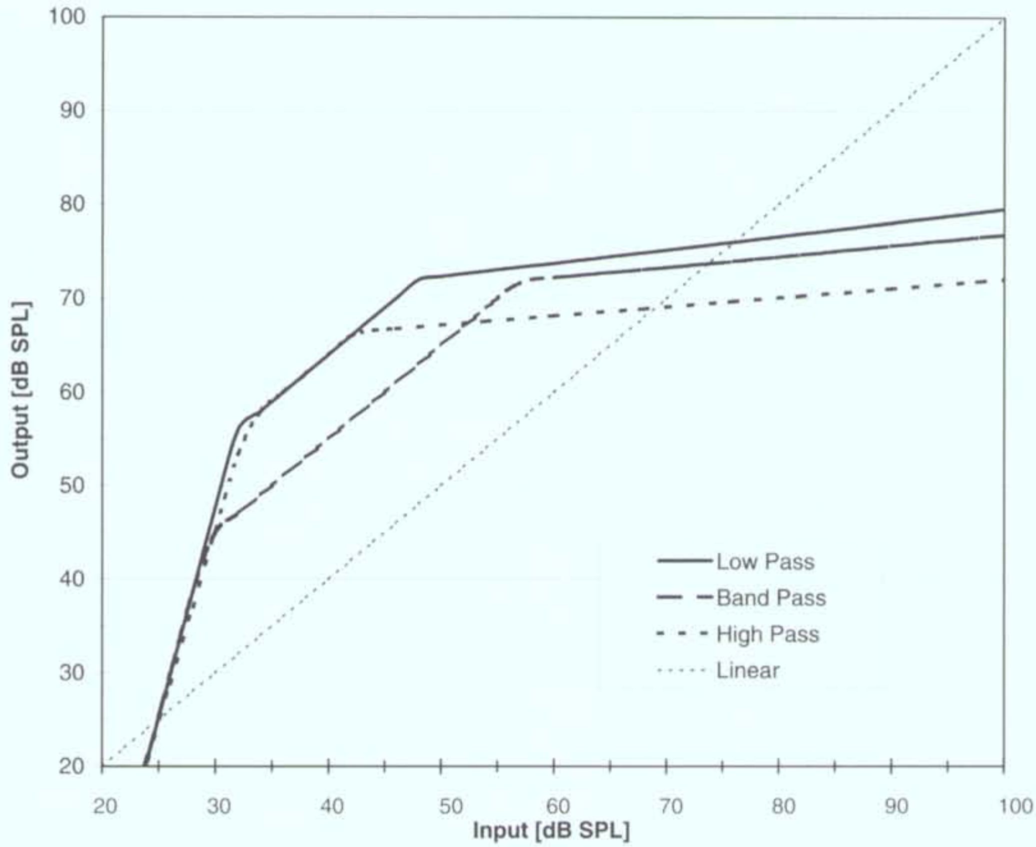


Fig. 4: Example of an I/O characteristic actually implemented for algorithm AV in the three frequency channels adjusted for one test subject (MW, right ear). The I/O characteristics for the low-pass and high-pass channel had to be limited to a maximum of 24 dB (maximum possible gain given by the implementation of the gain table). The maximum gain of the band-pass signal had to be further reduced to avoid feedback.

Abb. 4: Beispiel für die tatsächlich implementierten I/O-Charakteristiken für den AV-Algorithmus in den drei Frequenzkanälen eingestellt für einen Probanden (MW, rechtes Ohr). Die I/O-Charakteristik im Tief- und Hochpasskanal musste auf eine maximale Verstärkung von 24 dB begrenzt werden (maximal einstellbare Verstärkung in der Verstärkungstabelle). Die maximale Verstärkung im Bandpasskanal musste zur Vermeidung von Rückkoppelung zusätzlich begrenzt werden.

5.1 Sequence of measurements

For the audiological classiffication of the subjects, a set of tests was first performed for the unaided condition, which includes standard audiometry, categorical loudness scaling experiments and speech intelligibility tests in noise. In the subsequent first laboratory test, the DASi-2 prototype hearing-aid was fitted to the individual subjects hearing loss as described in section 4. A series of loudness scaling and speech intelligibility tests was then performed for the aided condition. After these measurements, subjects tested the device in everyday life for at least 14 days in order to get accustomed to the »new« hearing-aid. Subjects were told to use and compare all algorithms in all relevant listening environments. After this period, loudness scaling experiments with broadband signals and intelligibility tests were repeated in the laboratory. In addition, paired comparison quality tests were conducted and the subjects had to answer a questionnaire concerning their everyday experience with the hearing-aids in real life.

5.2 Category Loudness Scaling

Loudness perception was individually measured using a category loudness scaling procedure described by *Hohmann* and *Kollmeier* (1995b) and *Kollmeier* (1997b) with an 11-category scale (i.e., 5 main categories »very soft«, »soft«, »medium«,

Program number	Algorithm
A	*CL*, compression limiting
B	*SC*, syllabic compression
C	*AV*, automatic volume control
D	*NR* & *DC*, combination of the »best« noise reduction algorithm with the »best« dynamic compression scheme.

Table 2: Assignment of the DASi-2 program number and the hearing-aid configuration used for the field test.

Tab. 2: Zuordnung der Programmnummern des DASi-2 zu den im Rahmen des Feldtests verwendeten Verarbeitungsstrategien.

»loud« and »very loud«, plus 4 intermediate categories, between these as well as the two limiting categories »inaudible« and »too loud«). The procedure estimates the loudness given in categories as a function of the signal level and was used for estimating loudness perception of several stimuli for the unaided as well as for the aided condition. First of all, loudness scaling was performed in order to prescribe linear amplification and compression I/O characteristics. These unaided measurements used narrowband signals which were presented monaurally via headphones (Sennheiser HAD 200). All other loudness scaling experiments, including the fine tuning of the amplification, were carried out aided under free field conditions in a sound-proof booth using one loudspeaker (FAR CR10-S) situated 1 meter directly in front of the subject. For these measurements the ITE of the contralateral ear was switched off in order to achieve monaural testing. These experiments were carried out with several narrowband stimuli as well as with a broadband noise (speech-shaped noise).

a) Narrowband noise signals

The stimuli used for the narrowband loudness scaling were third octave-band noises centered at 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 4 and 6 kHz of 2 seconds' duration, including cosine ramps of 50 ms. These stimuli were presented monaurally via headphones to obtain the data for the parameter prescription of each of the respective algorithms.

To evaluate and verify the effect of the different dynamic compression schemes on loudness perception, loudness scaling with the narrowband noise signals centered at 0.5, 1.5 and 3 kHz was performed directly after fitting the processing schemes and was repeated at the end of the field test period.

b) Broadband noise signals

To evaluate the effect of the different dynamic compression algorithms on loudness perception of broadband signals, loudness scaling was performed with speech-shaped noise of 2 seconds' duration, taken from the Göttingen Sentence Test material (*Kollmeier* and *Wesselkamp* 1997). Scaling was performed directly after fitting the parameters and repeated at the end of the field test period.

5.3 Adaptive Sentence Test

An adaptive sentence test (Göttingen Sentence Test) was used to measure speech intelligibility in speech-shaped noise. The sentence test is described in detail in *Wesselkamp* et al. (1992), *Brand* and *Kollmeier* (1996) and *Kollmeier* and *Wesselkamp* (1997).

In this study, all speech intelligibility tests were performed under free field conditions. The noise level in the aided conditions was set to 65 dB SPL, which corresponds to the (signal-specific) average MCL of normal-hearing listeners.

Note that for each subject all algorithms were adjusted to give the same loudness for this signal at this level. The noise level in the unaided condition was set individually to the subjects' individual MCL, which was derived from the broadband loudness scaling in the unaided condition. For each condition to be measured, 20 sentences (two test lists) were used to determine the speech reception threshold (*SRT*, which is defined as the signal-to-noise ratio at which 50 % word score is obtained). During the adaptive tests, the speech level was varied, but the noise level was kept fixed.

5.4 Quality measurements

The purpose of the quality measurements was to assess preference across the processing schemes with respect to how quality in a specific acoustical environment subjectively assessed. The assessment was performed for speech in quiet (presentation levels 45 and 65 dB SPL), speech in cafeteria noise (presentation levels 45 and 65 dB SPL) and music (i.e., a segment of a popsong at presentation levels of 45, 65 and 80 dB SPL).

The subjectively perceived quality of the processing schemes was measured in a complete »paired« comparison experiment, i.e., each scheme was compared to each other. The subject was

Verbal judgement	Score
»A is much poorer than B«	2 points for B
»A is poorer than B«	1 point for B
»A and B are the same«	no points
»A is better than B«	1 point for A
»A is much better than B«	2 points for A

Table 3: Verbal categories for rating the difference in overall quality between two processing schemes. For further analysis, these categories were transformed into the numerical scores given in this table.

Tab. 2: Verbale Kategorien zur Beurteilung des Unterschiedes in der Gesamtqualität zweier Verarbeitungsstrategien (linke Spalte). Zur weiteren Analyse wurden den Urteilen Punkte zugeordnet (rechte Spalte).

asked to compare the »overall quality impression« for two schemes on a verbal scale. The answers were transformed into scores according to Table 3.

For each comparison, the two programs to be compared were stored at two neighbouring positions on the program switch of the DASi-2 prototype hearing-aid. The subject was allowed to listen to the test stimuli and to switch between the two programs at will.

5.5 Questionnaire and interview

After long-term testing, the subjects had to answer a questionnaire concerning their everyday experiences while using the processing schemes in real life. The questionnaire focused on speech intelligibility, sound quality and loudness. The subjects' answered questions on each processing scheme that had been tested in everyday life. Each question concerned a specific subjective sound impression (such as naturalness) and could be answered by choosing from 5 response alternatives ranging from a positive to a negative rating.

In addition to the questionnaire, the subjects were informally interviewed about their experiences with the hearing-aid algorithms during the field test. The diary which the subjects were instructed to keep for this purpose was used as a starting point for the interview.

6. Results and discussion

6.1 Loudness scaling

Figure 5 shows the results of the narrowband loudness scalings measured shortly after the fitting of the processing parameters to the hearing-impaired subjects. The mean values averaged across the five subjects (10 ears) for the equal-loudness levels in the categories »very soft«, »medium« and »very loud« are plotted. Average data for normal-hearing listeners taken from *Hohmann* and *Kollmeier* (1995b) is included as a reference. As expected from the fitting procedure, all algorithms do restore normal loudness at input levels corresponding to the normal MCL (loudness impression »medium« for normal-hearing subjects). However, the two processing schemes exhibiting a linear characteristic within the relevant range of input levels (*LIN* and *CL*) were not able to restore normal loudness at low input levels. The full range dynamic compression schemes (*SC* and *AV*), however, were able to restore normal loudness for low and medium input levels. It can be seen from Figure 5 that algorithms *SC* and *AV* yielded the same results in this experiment. Obviously, the effective compression for the signals presented was the same for both algorithms, i.e., the higher compression ratio of scheme *AV* is compensated for by the longer time constants. Hence, the aim of approximately equating the effective compression ratio for algorithms *SC* and *AV* was realized.

Loudness scaling measurements according to those shown in Figure 5 were repeated for algorithms *CL*, *SC* and *AV* at the end of the field test. These showed a consistent shift towards higher levels for all algorithms, i.e., the same signals were generally judged to be softer at the end of the field test than at the beginning of the field test. For low and medium levels, this shift was about 5 dB. It amounts to 10 dB for high levels. This is most probably an effect of acclimatization.

The broadband loudness scalings (Figure 6) show behaviour similar to the narrowband scalings. The prescription goal for algorithms *LIN* and *CL* is perfectly achieved, i.e., a match between normal-hearing and hearing-impaired MCL's. However, *LIN* and *CL* were only able to shift the hearing-impaired listeners dynamic range without extending it. In contrast, the *SC* and *AV* partially achieved their respective prescriptive objetives of extending the hearing-impaired listener's dynamic range. However, hearing-impaired listeners still perceived high levels as too loud compared to normal-hearing listeners. The repetition of the broadband loudness scalings at the end of the field test (dashed lines with

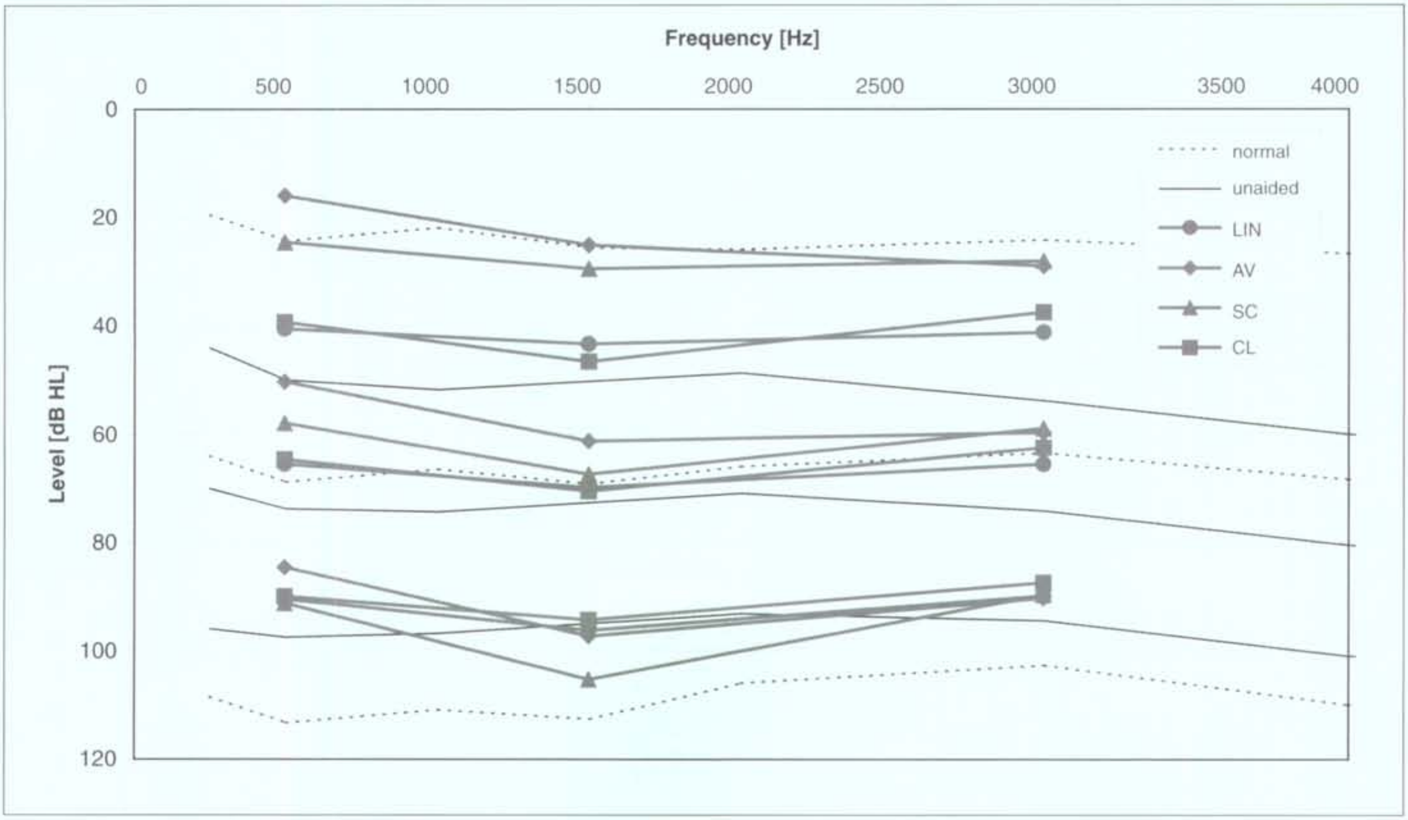


Fig. 5: Mean results of the narrowband loudness scalings with algorithms LIN , AV, SC and CL as well as in the unaided condition. Curves of equal loudness for the loudness categories »very soft«, »medium« and »very loud« are plotted, respectively (mean values over the five hearing-impaired subjects). The dotted lines show the respective normal-hearing data.

Abb. 5: Ergebnisse der Lautheitsskalierung mit schmalbandigen Signalen bei Versorgung mit Algorithmus LIN , AV , SC und CL, sowie unversorgt. Dargestellt sind die Kurven gleicher Lautheit für die Kategorien »sehr leise«, »mittelaut« und »sehr leise« (Mittelwerte über 5 schwerhörige Probanden). Die gepunkteten Kurven zeigen die Daten für Normalhörende.

open symbols in Figure 6) again showed an effect of acclimatization, especially at high levels. The data shows an increase in the subjects' dynamic range by 5 dB for algorithm *AV* and 10 dB for algorithm *SC*, respectively. It is not quite clear why this slight increase in the acclimatization effect occurs for algorithm *SC*.

6.2 Adaptive Sentence Test

The speech intelligibility measurements were carried out in speech simulating noise at a presentation level corresponding to the subjects' individual MCL's. The presentation level was therefore set to 65 dB SPL in the aided conditions and adjusted to the subjects individual MCL's (derived from the broadband loudness scaling) in the unaided condition. Hence, speech intelligibility was tested at the same overall loudness. The individual results are shown in Figure 7. Due to high inter-individual variability, mean and median values shall not be discussed here. Differences in SRT of less than about 1 dB are not considered significant. Except for subject HM, all test subjects performed better unaided (i.e., without any change in the frequency spectrum and dynamic range) than aided with all of the algorithms. Because the test stimuli were presented at the same overall loudness and because of the comparatively flat hearing loss of most subjects, this was more or less to be expected because the compression and the linear frequency shaping in the hearing-aid does not provide any additional benefit. On the contrary, the occlusion of the ear, the restricted receiver transmission quality (limited frequency range, nonlinear distortion) as well as hearing-aid noise (originating from the microphone, analogue circuitry and quantization noise) will limit the performance with a hearing-aid considerably. Hence, speech intelligibility was best without a hearing-aid at the appropriate input levels. Only subject HM shows a benefit in speech intelligibility in the aided conditions.

For these reasons, the processing schemes were compared to each other and not to the unaided condition. Figure 8 shows the

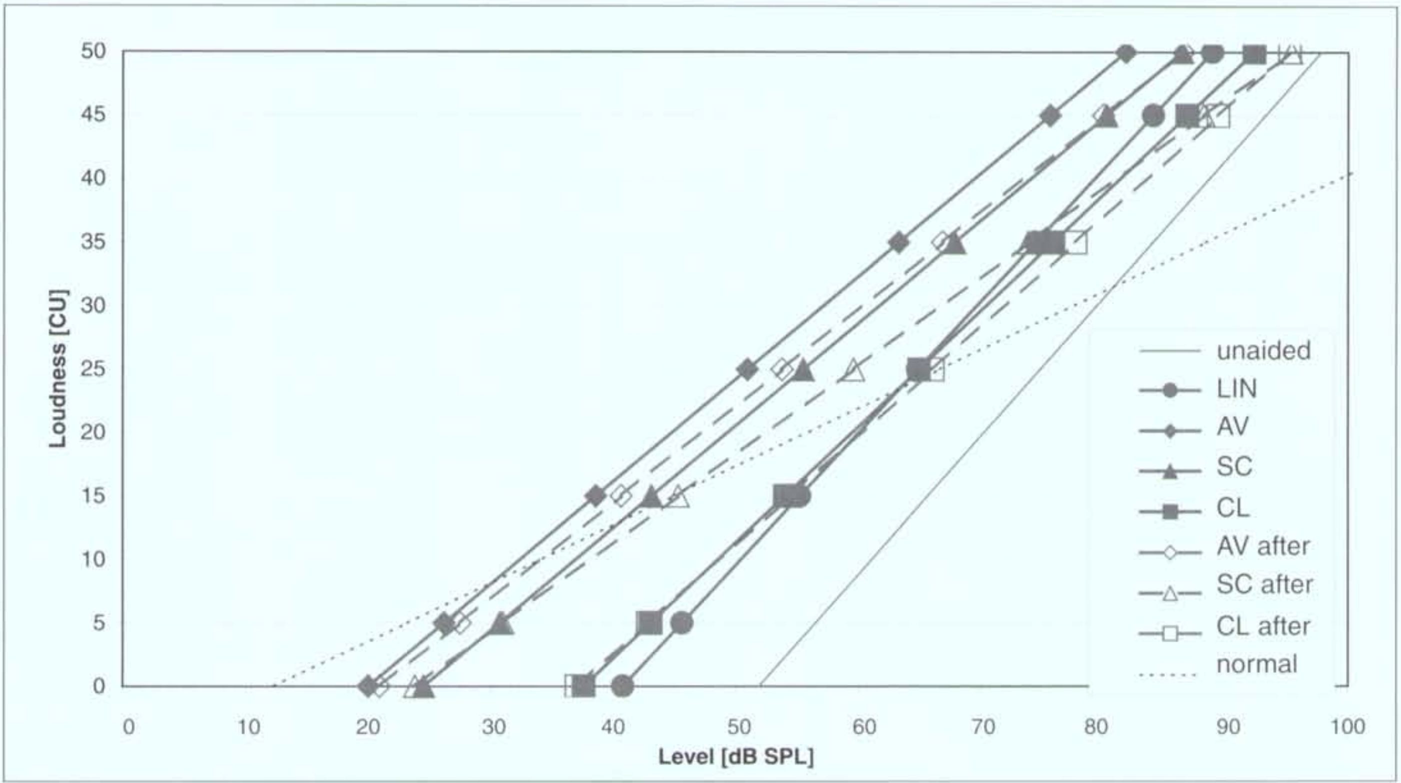


Fig. 6: Results of the broadband loudness scalings with algorithms LIN , AV, SC and CL as well as unaided. Each curve represents the mean of the individual linear functions fitted to the individual measurement points of the four hearing-impaired subjects before (solid lines, filled symbols) and after (dashed lines, open symbols) the field test. Symbols are shown for 0 CU = »not heard«, 5 CU = »very soft«, 25 CU = »medium«, 45 CU = »very loud« and 50 CU = »too loud«. The dotted lines show the respected normal-hearing data.

Abb. 6: Ergebnisse der Lautheitsskalierung eines breitbandigen Signals bei Versorgung mit Algorithmus LIN , AV, SC und CL, sowie unversorgt. Die Kurven zeigen die Mittelwerte über die individuellen Lautheitsfunktionen gemessen bei vier Schwerhörenden vor (durchgezogene Kurven, gefüllte Symbole) und nach dem Feldtest (gestrichelte Kurven, offene Symbole). Die Symbole entsprechen den Kategorien 0 CU = »nicht gehört«, 5 CU = »sehr leise«, 25 CU = »mittellaut«, 45 CU = »sehr laut« und 50 CU = »zu laut«. Die gepunkteten Kurven zeigen die Daten für Normalhörende.

data from Figure 7 for schemes *AV* and *SC* relative to scheme *LIN*. No clear trend could be found in the performance of schemes *AV* and *SC* as compared to the reference. Subject BD performed best with *LIN* , subject MW performed best with *AV* and the performance of all other subjects was the same for all of the processing schemes. This was more or less to be expected because all the schemes were fitted to give the same amplification for the 65 dB SPL speech simulating noise presented in this experiment. It can be concluded that the dynamic compression algorithms neither improve nor deteriorate speech intelligibility in noise at medium levels as compared to linear processing. On the other hand, at low signal levels, the dynamic compression algorithms will certainly outperform algorithm LIN simply because of audibility.

6.3 Quality measurements

The perceived quality of the processing schemes was again measured in a »paired« comparison experiment. Quality scores were calculated as the sum of the numerical values according to Table 3 across all paired comparisons for each condition, respectively. The results are given in figures 9 to 11. The median values are shown together with the interdecile ranges I_{80}, which are a dispersion measure covering 80 % of the distribution of scores. Because each of the four processing schemes was compared to each other, a maximum score of 6 could be achieved for a scheme

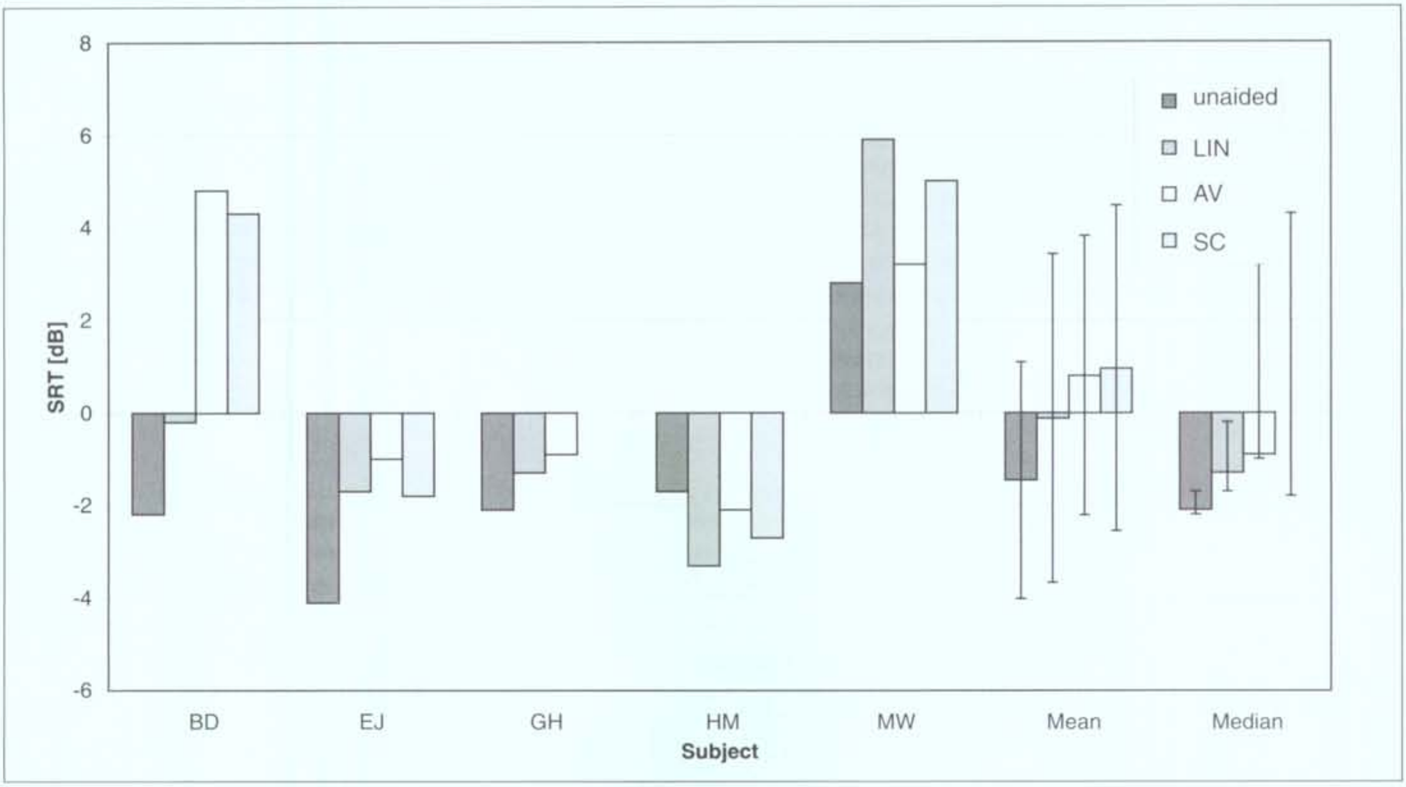


Fig. 7: Results of the speech intelligibility measurements in noise with processing schemes LIN , AV and SC as well as in the unaided condition. The speech reception thresholds (SRT) are given for each subject and for each algorithm. The two right columns show the mean and median data averaged across subjects, error bars denote plus or minus one standard deviation and the interquartile range, respectively. Lower values indicate better performance.

Abb. 7: Ergebnisse der Sprachverständlichkeitsmessungen in Störgeräusch bei Versorgung mit Algorithmus LIN , AV , SC und CL, sowie unversorgt. Dargestellt sind der Signalrauschabstand der zu einem 50%-igen Sprachverstehen führt (speech reception thresholds, SRT) für jeden Probanden und jeden Algorithmus. Die beiden rechten Spalten zeigen die Mittelwerte bzw. Mediane über die Versuchspersonen mit Standardabweichung bzw. Interquartilsbereich. Niedrigere Werte kennzeichnen bessere Verständlichkeitsschwellen.

that was judged as »much better« in all comparisons (3 comparisons * 2 scores per comparison), whereas a score of zero would indicate that this scheme was never judged better in quality as compared to any other scheme.

Figure 9 shows the results sorted by presentation level and averaged across subjects and test stimuli. It can be seen that the dynamic compression schemes (*AV* and *SC*) get higher scores compared to the linear schemes (*LIN* and *CL*) at low and medium input levels. This holds true particularly for low levels and also for *AV* whereas the differences between the schemes decrease as levels increase. Only one condition was tested at high levels (»music« at 80 dB SPL) and no clear difference between the algorithms was observed.

A more differentiated view of the results is provided by Figure 10, which shows the results for each of the test stimuli »speech in quiet«, »speech in cafeteria noise« and »music« averaged across subjects and presentation levels. Here *SC* performaned well under all conditions, whereas *AV* achieved very good scores for »speech in quiet« and »music« at low levels, but its performance decreased for stimuli that employ impulsive noise bursts at higher levels (»speech in cafeteria noise«). Focusing on the linear schemes *LIN* and *CL*, it can be seen that their overall performance is poorer compared to the dynamic compression schemes. An exception here is »music« at high levels, where *LIN* gets the highest score. The reason for this might be the subjects' acceptance of high levels in this specific situation and also that *LIN* provided the highest output level in this situation. Another exception to the poorer performance of the linear schemes in gen-

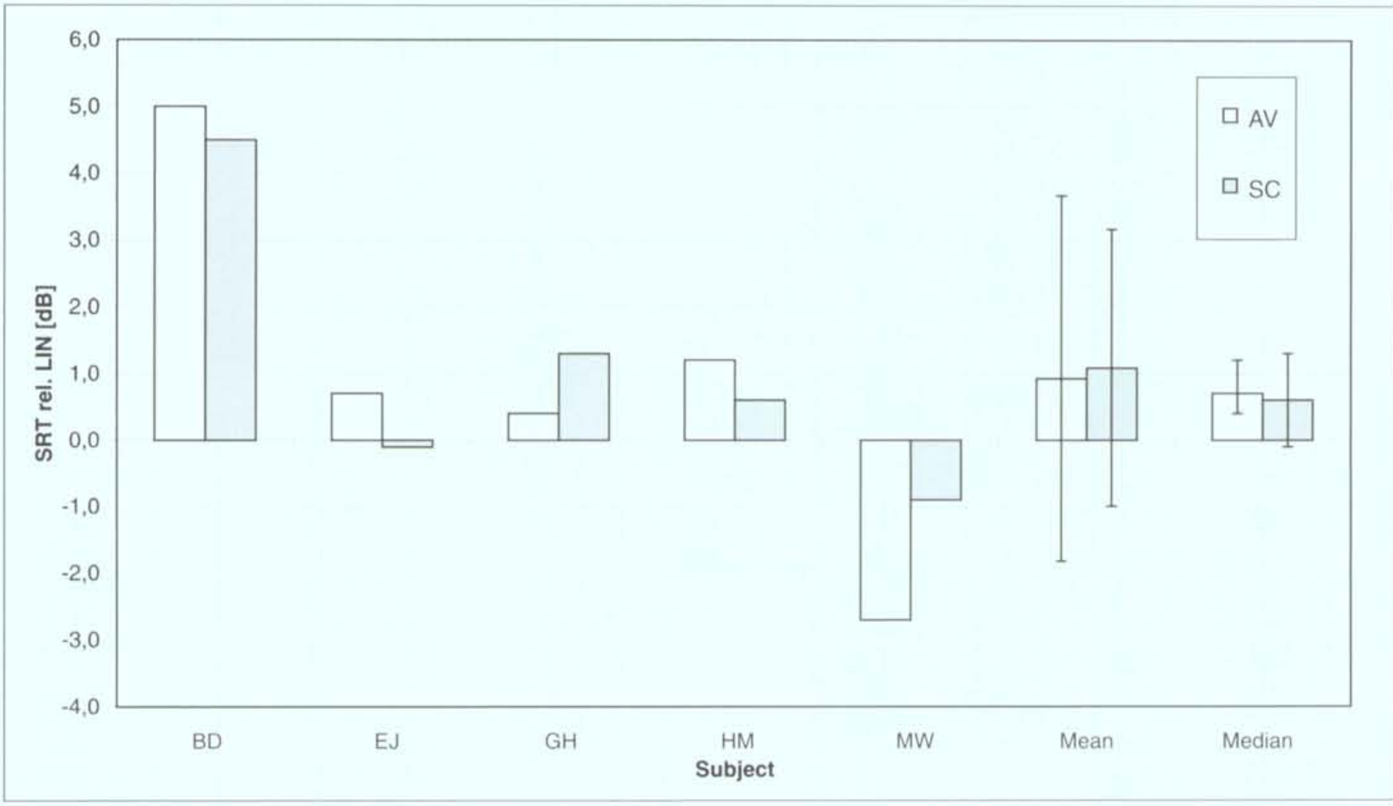


Fig. 8: Individual SRTs for algorithms AV and SC plotted relative to the SRT observed with algorithm LIN. Values were calculated from the data shown in Figure 7.

Abb. 8: Individuelle Sprachverständlichkeitsschwellen (SRT) für die Algorithmen AV und SC relativ zu dem mit dem Algorithmus LIN gemessenen SRT. Die Werte wurden aus den in Abbildung 7 gezeigten Daten berechnet.

eral is *CL* in the »speech in cafeteria noise« situation at the highest level. Here, the ability of *CL* to effectively limit high-level peaky sounds seems to be advantageous for some subjects.

In order to analyze individual differences in the quality judgements, Figure 11 shows the results for each subject averaged across presentation levels and test stimuli. It can be seen that subjects BD, EJ and HM did not show a clear preference for any of the processing schemes, whereas subject GH showed a clear preference for AV, and subject MW for both dynamic compression schemes (*AV* and *SC*). The latter may have been influenced by the restrictions which had to be made with regard to the I/O characteristic of algorithm *AV* for subject MW, i.e., subject MW might have given higher scores for algorithm *AV* if the I/O characteristic could have been fitted exactly to its prescription (see Figure 4). The individual results for subjects GH and MW can be explained by the fact that their residual dynamic range was the smallest within the group of subjects (c.f., Table 1). The effect of compression was therefore expected to be most pronounced in these subjects.

When all the data for all stimuli, levels and subjects (Figure 11, bars on the right) is considered, only slight differences between the algorithms could be determined.

After having tested the different processing schemes in real life conditions, the subjects filled out the questionnaire described in section 5.5. To evaluate the results, the judgements were transposed from the verbal scale into scores varying from 0 to 4, where 0 was the most negative and 4 the most positive rating of the algorithm in each question. The results are given in Figures 12 to 15. Subjective speech intelligibility was generally judged to be better in the dynamic compression schemes (*AV* and *SC*) than in scheme *CL* (Figure 12). The best results were obtained for *AV*. This is found consistently to all listening conditions. Generally speaking, the rating for »speech in noise« decreased slightly for all schemes compared to the other situations. This can be explained by the principle difficulty of the hearing-impaired to understand speech in background noise.

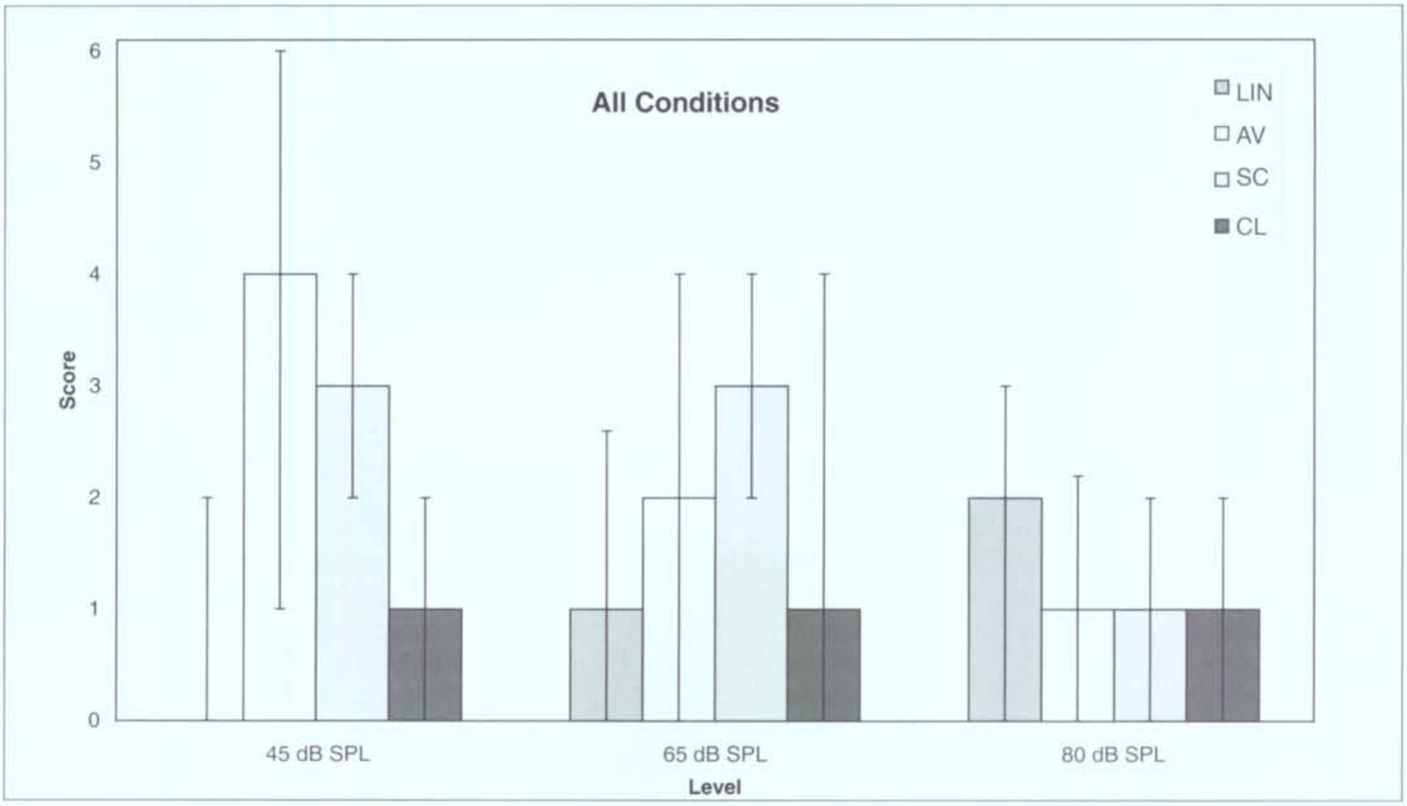


Fig. 9: Results of the paired comparison test sorted by presentation level and averaged across subjects and test stimuli. The median scores with the interdecile range I_{80} for processing schemes LIN , AV, SC and CL are given for each presentation level: 45 dB SPL (15 observations), 65 dB SPL (15 observations) and 80 dB SPL (5 observations).

Abb. 9: Ergebnisse der Paarvergleiche sortiert nach Darbietungspegel und gemittelt über die Probanden und Teststimuli. Dargestellt sind die Mediane und Interdezilbereiche I_{80}.

The judgements concerning sound quality (Figure 13) show no clear preference for anyone of the processing schemes, but there is a tendency for higher ratings to be given to the dynamic compression schemes, especially to *AV*. However, the data shows that CL performed well in the ratings for sound naturalness. Thus, it can be assumed that linear processing comes closest to the subjects' impression of how their acoustical environment normally sounds. But the better results for *AV* in sound clarity show that »subjectively« the processing scheme providing the most natural sound is not automatically the scheme that gives the highest sound clarity, or indeed, the best speech intelligibility, as Figure 12 shows. Figure 14 gives the median subjective loudness judgements from the questionnaire for all subjects and processing schemes. It reveals that the loudness judgements at the end of the field test agree with the results of the loudness scaling (see 6.1). Algorithm *CL* was judged rather too soft, *AV* rather too loud, whereas *SC* was judged as comfortable. On the other hand, there are clear differences in the individual judgements. Algorithm *AV* provided the correct amplification for subjects GH and MW, whereas *CL* was best for subjects BD and HM. As argued in section 6.3, the preference of subjects GH and MW for processing scheme *AV* can be explained by the fact that their residual dynamic range was the smallest within the group of subjects. As can be observed from their individual loudness scaling data, algorithm *AV* is best able to widen their dynamic range to the dynamic range of normal hearing.

These individual differences were also found when the median across all speech and sound quality judgements was calculated for each subject individually[4] . Again, algorithm *AV* was best for subjects GH and MW, whereas *CL* was best for subjects BD and might be best for subject HM.

All in all, the results of the questionnaire assessment suggest that scheme *SC* is a good compromise between schemes *CL* and *AV* (Figures 12 and 13). But the individual data shown in Figures 14 and 15 reveal that there are large inter-individual differences between the subjects. It can be concluded that some subjects in general prefer algorithm *AV*, whereas other subjects prefer *CL*. This finding is supported by the appearance of larger error bars

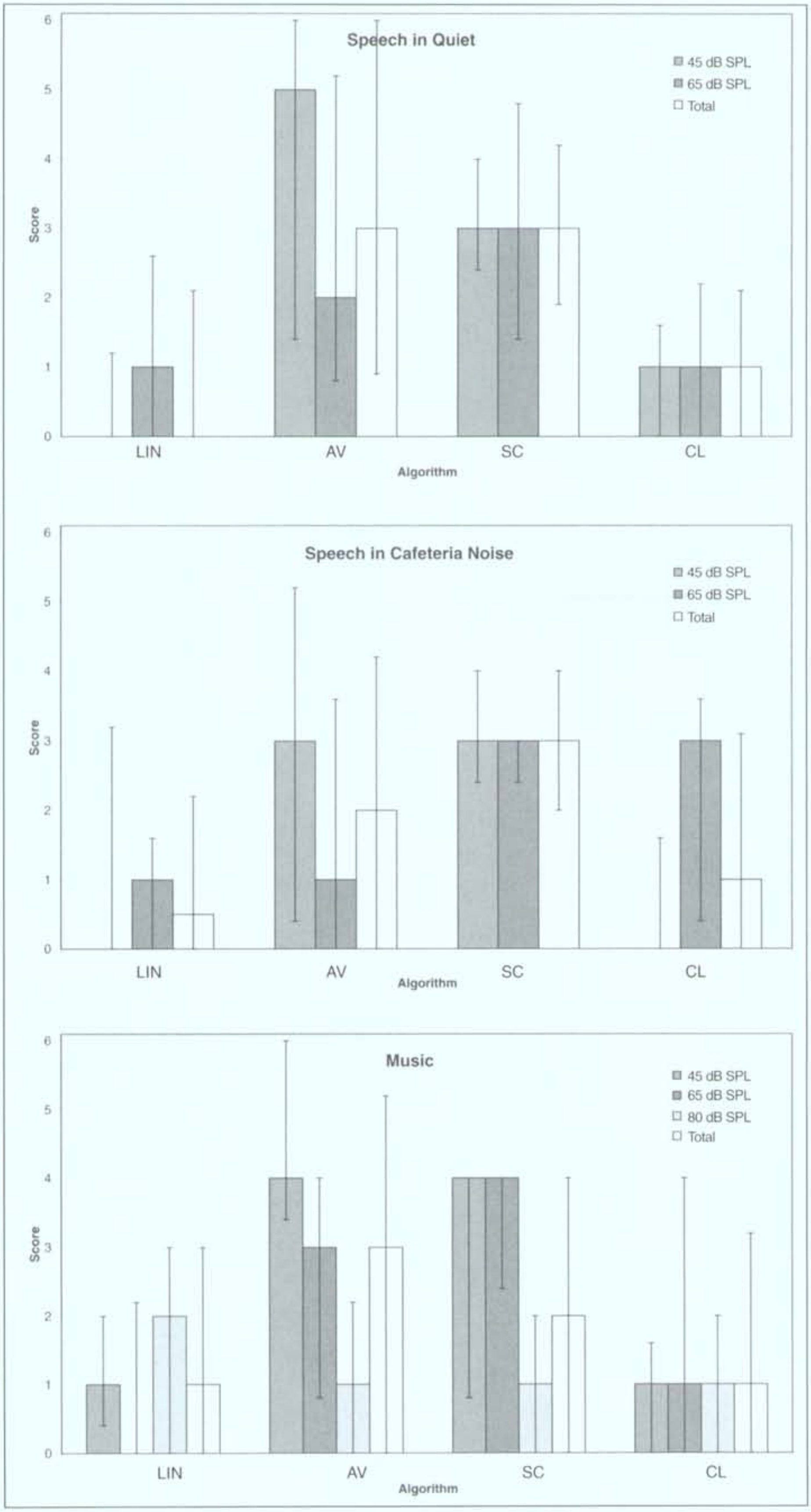


Fig. 10: Results of the paired comparison test for »speech in quiet« (top), »speech in cafeteria noise« (middle) and »music« (bottom). The graphics show the median scores with interdecile range I80 across the 5 subjects for the processing schemes LIN, AV, SC and CL, respectively, subdivided into the presentation levels. For each scheme the far right bar denotes the median across all subjects and all presentation levels.

Abb. 10: Ergebnisse der Qualitätstests im Paarvergleich in den Situationen »Sprache in Ruhe« (oben), »Sprache in Cafeteriageräusch« (mitte) und »Musik« (unten). Dargestellt sind die Mediane der erzielten Punktzahlen und der Interdezilbereich I80 gemittelt über 5 Probanden für die Verarbeitungen LIN, AV, SC und CL (von links nach rechts), jeweils unterteilt in die verschiedenen Darbietungspegel, sowie jeweils rechts die entsprechenden Gesamtmittelwerte über die Probanden und Darbietungspegel.

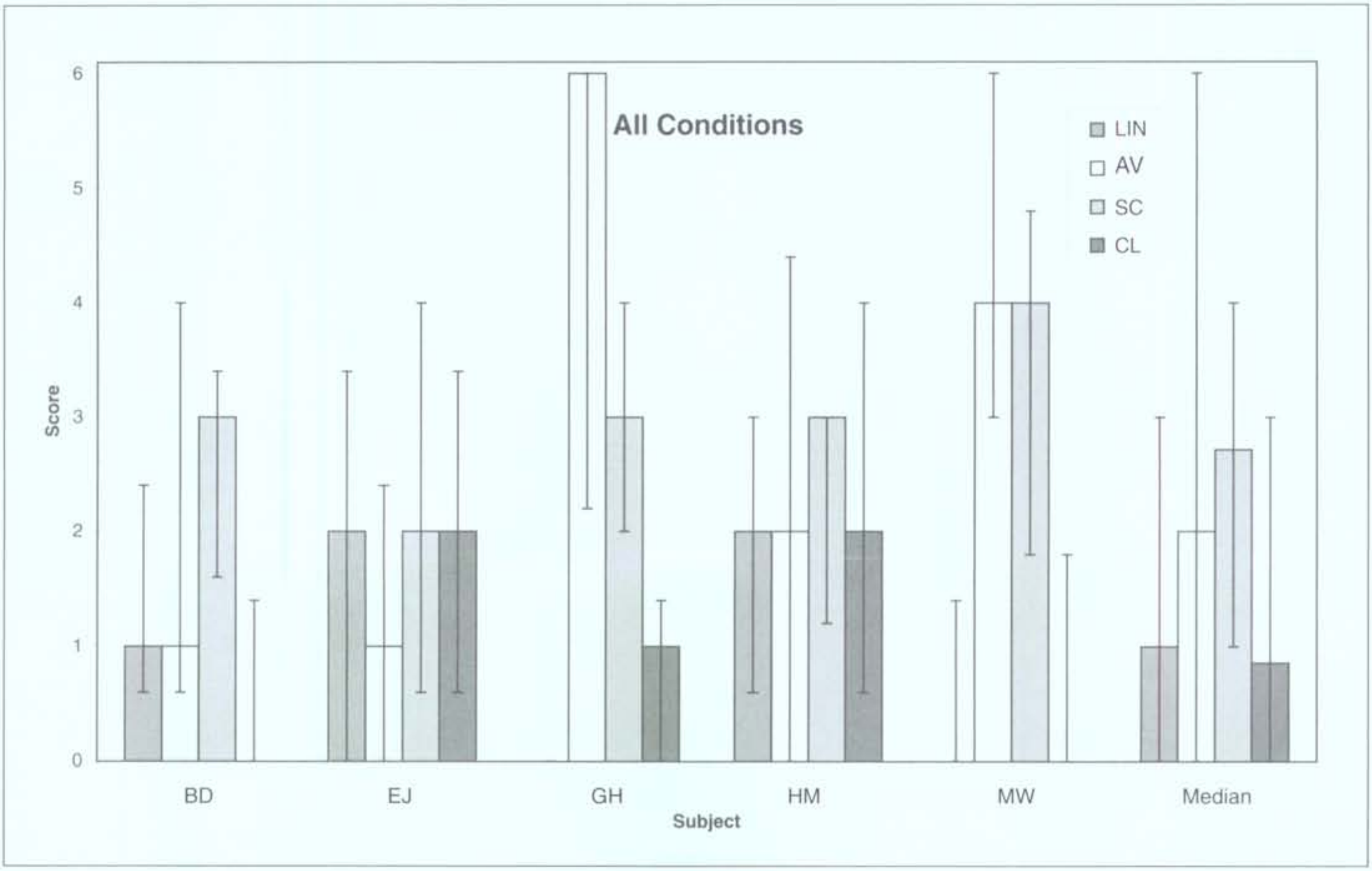


Fig. 11: Individual results of the paired comparison test, showing the median scores with interdecile range I_{80} averaged across all stimuli and level combinations for processing schemes LIN, AV, SC and CL (7 observations per algorithm). The far right bars denote the median over stimuli, levels and subjects for each algorithm (35 observations per algorithm).

Abb. 11: Individuelle Ergebnisse der Paarvergleiche. Dargestellt sind die Mediane der erzielten Punktzahlen und der Interdezilbereich I_{80} gemittelt über alle Stimuli und Darbietungspegel (7 Observationen pro Algorithmus). Die rechte Spalte zeigt die Mediane über alle Kombinationen aus Stimulus, Pegel und Proband (35 Observationen pro Algorithmus).

in the results for algorithm *AV* and *CL* as compared to the error bars for algorithm *SC*.

6.5 Subjective Assessment (Interview)

The following summarizes some of the results of the interviews with the test subjects after the field trial period. Most subjects reported that they had no problems handling the device, but complained about the inconvenience caused by wearing the DASi-2 prototype hearing-aids. The cable connections between the ITE devices and the speech processor were particularly disliked.
In accordance with the results given so far, subject BD liked scheme *CL* best and found that scheme *AV* was too loud and had too many background noises. Scheme *SC* was better for unintelligible speech on television. Subject HM did not like the DASi-2 prototype hearing-aids at all. Of all the schemes, HM liked *CL* best because other people's voices were most intelligible. Like subject BD, she complained about the loud background noises produced by scheme *AV*. In addition, she was annoyed about the tendency to feedback exhibited by scheme *AV*.

Subjects GH and MW on the other hand, remarked that scheme *CL* was much too soft for most real life situations. Both subjects liked *AV* best and preferred using it at the theatre and cinema in

[4] It is critical to calculate the median across data for questions pointing to quite different attributes and which are also answered on different verbal scales. But here we just wanted to clarify if there were inter-individual differences in the judgements which are consistently among all these attributes.

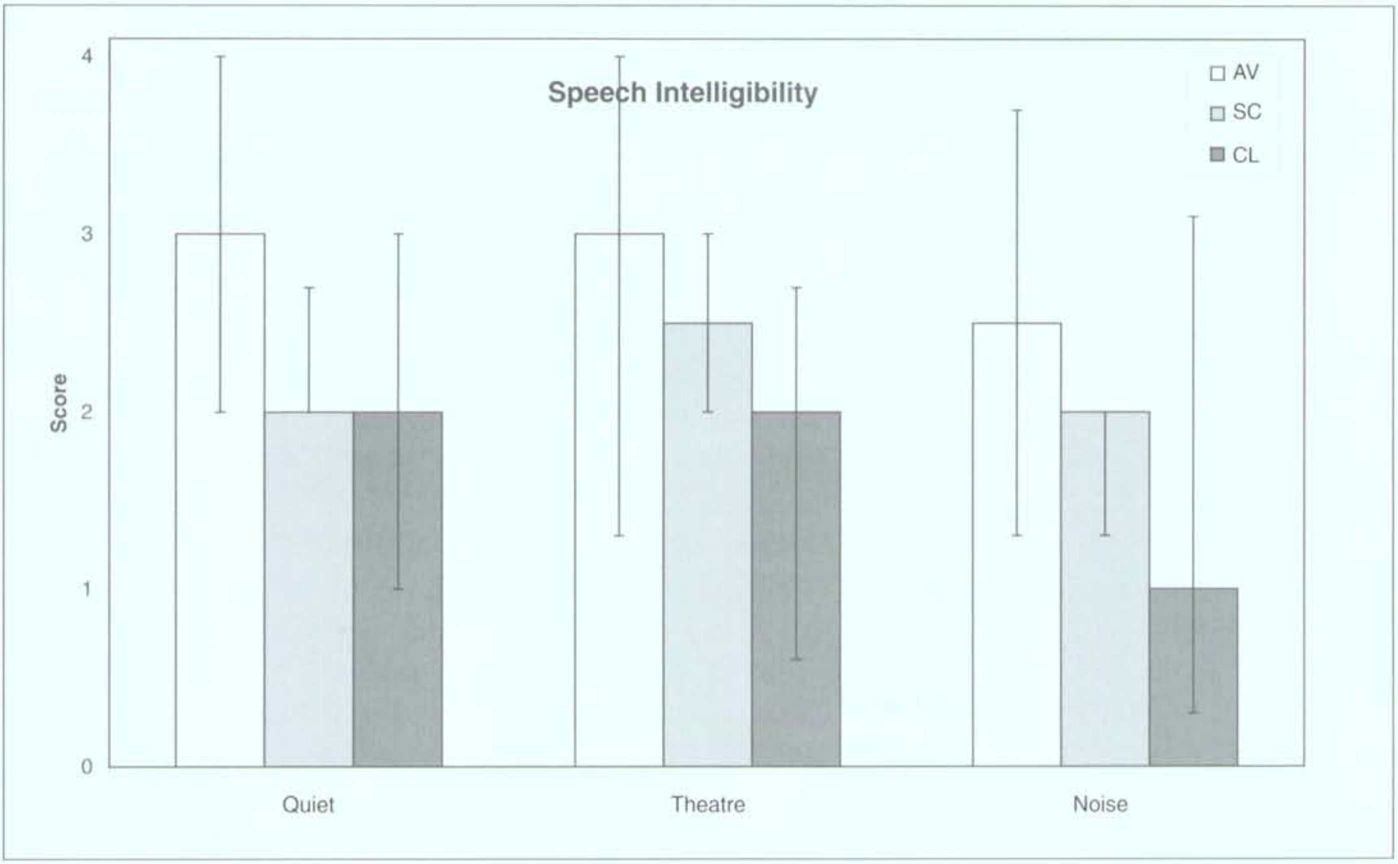


Fig. 12: Subjective judgements concerning speech intelligibility in real life conditions regarding the criteria »Speech in quiet«, »Speech in theater or lecture« and »Speech in noise« (questionnaire assessment). The median scores across subjects for the processing schemes AV, SC and CL are plotted for each situation, together with their interdecile range I_{80}.

Abb. 12: Subjektive Beurteilung der Sprachverständlichkeit in Alltagssituationen anhand der Beurteilungen im Fragebogen unterteilt in »Sprache in Ruhe«, »Sprache in Theater oder Vorlesungen« und »Sprache in Störgeräusch«. Dargestellt sind die Mediane über alle Probanden für die Verarbeitungen AV , SC and CL, sowie deren Interdezilbereiche I_{80}.

particular. Subject GH praised the excellent sound quality of *AV*, which was quite similar to his own hearing-aids. Subject MW preferred *AV* because it gave the best speech intelligibility. MW voted for *SC* as giving the best sound quality because its sound was the softest.

7. Summary and Discussion

A battery of tests was performed with five hearing-impaired subjects fitted with a prototype digital hearing-aid in order to compare different 3-channel dynamic compression algorithms. Due to the carefully selected control conditions (i.e., unaided situation at roughly the same presentation level as in the aided situations, same frequency response for all algorithms at a medium input speech level with a speech-shaped input signal), the differences across algorithms were only very slight. Hence, no »winner algorithm« could be derived from the data. However, the controlled experimental design (battery of tests with a set of algorithms fitted with the same fitting rationale) provides information for a variety of aspects that may be relevant for other dynamic compression schemes as well:

Loudness scaling

Loudness scaling data with narrowband stimuli revealed that all algorithms achieved their main fitting goal, i.e., approximately restoring the loudness contour »medium« for narrowband signals across different frequencies. For low input levels, the linear algorithms (including the compression limiting algorithm) due to their construction principles were not able to provide sufficient amplification. For high input levels, on the other hand, all algorithms provided too much amplification, even though the

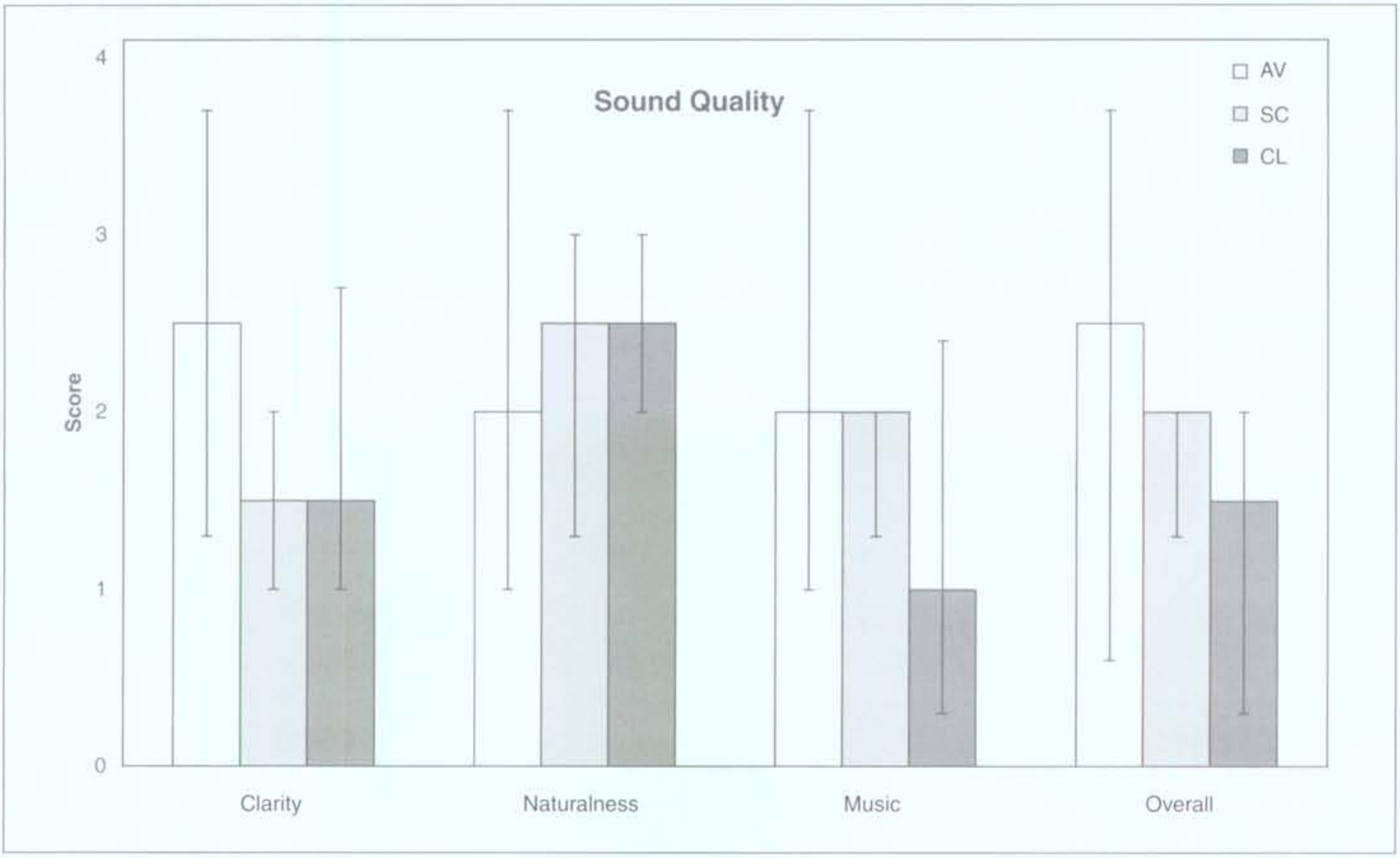


Fig. 13: Subjective judgements concerning sound quality in real life conditions regarding the distinguished criteria »naturalness of sound«, »clarity of sound«, »sounding of music« and »overall sound quality« (questionnaire assessment). For each situation the median scores across subjects for the processing schemes AV, SC and CL are plotted together with their interdecile range I_{80}.

Abb. 13: Subjektive Beurteilung der Verarbeitungsqualität in Alltagssituationen anhand der Beurteilungen im Fragebogen unterteilt in »Natürlichkeit des Klanges«, »Klarheit«, »Klang von Musik« und »Gesamt Klangqualität«. Dargestellt sind die Mediane über alle Probanden für die Verarbeitungen AV, SC and CL, sowie deren Interdezilbereiche I_{80}.

compression algorithms should in principle exhibit less amplification here than the linear algorithms. A similar finding can be derived for the broadband loudness scaling (cf. Figure 6): While the algorithms provide their correct gain at medium levels, they do not on average provide enough compression, i.e., they show too little amplification at low input levels and too high an amplification at high input levels. Although the compression algorithms perform better in this respect (by providing an enlarged input dynamic range of approximately 15 dB immediately after fitting and 5–10 dB more after acclimatization), they provide too much amplification at medium levels and do not provide enough dynamic compression to map the whole dynamic range of normal listeners into the remaining dynamic range of the hearing-impaired patients considered here. One reason for the insufficient amplification of the dynamic compression algorithms at low levels are the feedback problems encountered in real-time processing at high insertion gain values, which limit the maximum gain achievable by the wearable device. In this respect, the study with the wearable device differs considerably from the original experiments performed by *Appell* et al. (1995) using the same dynamic compression algorithms on a laboratory computer setup.

As already noted above, a certain acclimatization effect (according to *Gatehouse* 1992) was observed for the compression algorithms after 4 to 6 weeks of using the wearable prototype hearing-aid in daily life. This yields a 5–10 dB extension of the input dynamic range. However, it is not clear why the syllabic compression algorithm showed a higher acclimatization effect than the automatic volume control algorithms (*AV*), even though the difference between them is comparatively slight (5 dB).

Fig. 14: Subjective judgements concerning loudness in real life conditions (questionnaire assessment). Loudness values -2, 1, 0, 1 and 2 correspond to the loudness impression »much too soft«, »too soft«, »ok«, »too loud« and »much too loud«, respectively. The far right bars denote the median across all subjects.

Abb. 14: Subjektive Beurteilung der Lautstärke in Alltagssituationen anhand der Beurteilungen im Fragebogen. Den Werten -2, 1, 0, 1 und 2 entsprechen den Lautheitseindrücken »viel zu leise«, »zu leise«, »ok«, »zu laut« und »viel zu laut«. Ganz rechts dargestellt sind die entsprechenden Mediane über alle Probanden, sowie deren Interdezilbereiche I_{80}.

Sentence test

In general, no improvement in sentence intelligibility in noise could be observed for any of the processing schemes implemented on the wearable device (including the linear amplification). This finding is in line with most of the reports found in the literature (*Festen* 1999; *Hohmann* and *Kollmeier* 1995a; *Lunner* et al., 1998; *Stone* et al. 1997; *Verschuure* et al. 1998; *Walker* and *Byrne* 1984), including laboratory experiments. It also contradicts a few reports about intelligibility improvements for certain dynamic range compression algorithms (*Benson* et al. 1992; *Moore* et al. 1999).

The main reason for this lack of measurable benefit is the carefully selected reference condition that provides some overall level adjustment to compensate for the »attenuation« component of the hearing loss (see Results section). Even comparing the implemented algorithms to each other shows no systematic improvement or deterioration in the case of the dynamic compression algorithm over the linear amplification. This is probably due to the failure of the compression algorithms to provide more audibility for signal components as compared to the linear condition. This contrasts with some of the studies found in the literature that report a positive effect on speech reception thresholds both in quiet and in noisy conditions. These might have been a greater difference at lower presentation levels (where audibility limits the maximum performance both in quiet and, to a lesser degree, in noise) and at comparatively high presentation levels (where distortions introduced by the hearing-aid system and by peak clipping processing may limit the maximum intelligibility). However, the amount of intelligibility measurement blocks was limited by the number of available test sentences. Hence, no systematic intelligibility evaluation across a large dynamic speech

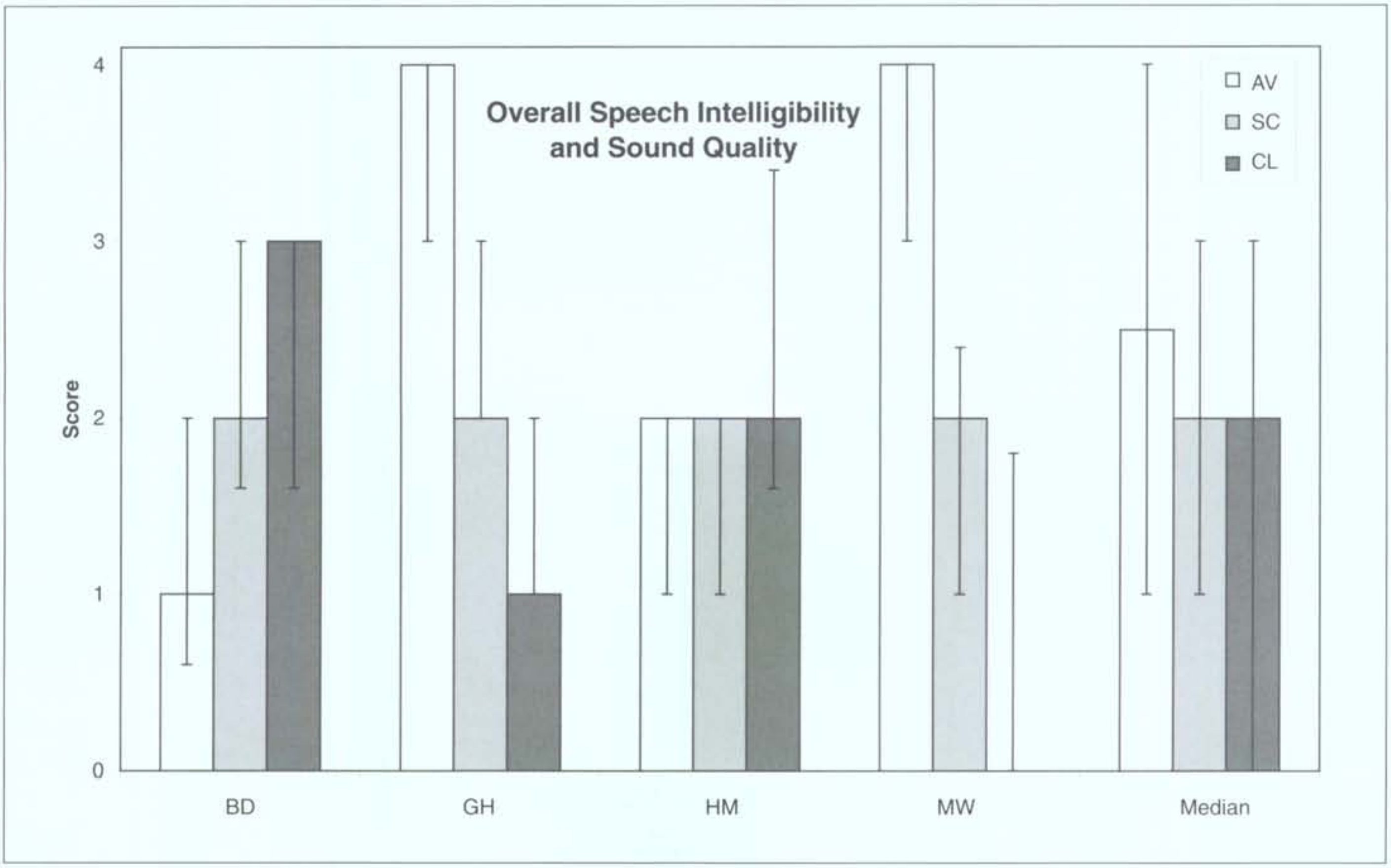


Fig. 15: Median values across all speech intelligibility and sound quality judgements for each subject and each processing scheme (questionnaire assessment). The far right bars denote the median across all subjects.

Abb. 15: Mediane gemittelt über alle Sprachverständlichkeits- und Klangqualitätsurteile für alle Probanden und alle Verarbeitungen. Ganz rechts dargestellt sind die entsprechenden Mediane über alle Probanden, sowie deren Interdezilbereiche I_{80}.

range was possible within the current study. This is in contrast to the previous study by *Appell* et al. (1995), where a wider range of test levels was employed and algorithm *AV* performed better than the others in general.

An unavoidable effect of the current experiments is that both the target speech and the interfering noise signal control the dynamic compression circuit. Hence, the detrimental effect of the noise may even be enhanced in the case of low signal-to-noise ratios (SNRs), since the overall gain may fluctuate synchronously with the noise fluctuations, whereas this effect is negligible at high (positive) SNRs. As a consequence – already noted by *Verschuure* et al. (1998) – the SNR corresponding to the unaided speech reception threshold (SRT) of the individual subjects influences the amount of this detrimental effect and may hence directly influence the potential benefit obtained from a dynamic compression algorithm in speech intelligibility tests. This may be the reason that subject MW (who showed the highest unaided SRT, the smallest residual dynamic range and required the highest compression) performed systematically better with the compression algorithms than the compared to algorithm *LIN*.

Quality judgements

The paired comparison test performed with various acoustic materials at three different levels clearly shows that compression (especially with algorithm *AV*) is preferable for low input levels, while syllabic compression seems to be advantageous at medium levels. No differences were found at higher levels. Of course, all of these algorithms (especially the linear algorithm) would need a peak clipping or compression limiting algorithm to prevent the user from output signals or sound peaks that are too high.

When the quality judgements are compared across subjects, it is striking that the subjects with the smallest dynamic ranges (i.e., subjects GH and MW) profited most from the dynamic compression algorithms. They gave algorithm *AV* the best overall scores, although it reacts comparatively slowly and hence cannot protect the user against sudden loud sounds very effectively. The results of the sound quality judgements (obtained by paired comparisons) are consistent with those of the questionnaire, where again *AV* performed best for speech (but no clear preference is given with respect to sound quality). On the other hand, the questionnaire results for loudness,were similar to those of the loudness scaling data: While algorithm *CL* was judged as rather too soft, *AV* was judged as a bit too loud. Again, there is very little variation, especially as regards the large individual differences across subjects shown by the error bars (cf. Figure 14). These individual differences are at least partly due to the differences in the residual dynamic range: as shown above, subjects with the smallest residual dynamic range (GH and MW) judged algorithm *AV* to be the best, while the other subjects preferred the compression limiting algorithm. The results of the informal interviews in general verify the data from the other experiments.

Generally speaking, the results of the quality comparison show that high amplification at low input levels (as provided by algorithm *AV*) is advantageous for hearing-aids that reflect the real world. They should therefore be combined with a strategy that ist able to better compensate for individual loudness deficits at medium and high levels by using a compression strategy such as syllabic compression. In addition, an effective reduction of sudden noise bursts at higher levels (provided by the compression limiting system) seems to be advantageous. Patients with a greatly reduced dynamic range appear to profit particularly from dynamic compression and judge these algorithms as capable of providing a higher signal quality on than linear amplification.

Although the general approach and fitting rationale behind the algorithms were meant to compensate for abnormal loudness perception in hearing-impaired listeners (at least within an intermediate level range), loudness compensation (perhaps most thoroughly provided by algorithm *SC*) does not necessarily lead to a higher speech intelligibility when compared to linear amplification and to automatic volume control at a certain speech level. This may well be due to our limited knowledge of loudness perception in hearing-impaired listeners and its relation to speech intelligibility in quiet and noise, which calls for further research in this area. Moreover, the field test results obtained here cannot be compared directly to laboratory tests due to the many restrictions of a »real-world« hearing-aid (such as limited feedback margin, limited frequency range and distortions and noise introduced by the hearing-aid setup). However, since the ultimate goal of dynamic compression algorithms is to provide user benefit in wearable hearing-aids for the real world, it would appear that field tests on such wearable devices shout be used for testing the performance of new algorithms.

Even though this study failed to produce clear-cut results, it still provides interesting general findings that could be used in design of future hearing-aids:

- No-one expressed a clear-cut preference for a specific dynamic compression algorithm over its competitors at medium input levels provided that some basic requirements are met (i.e., match of overall frequency shape and overall gain at intermediate levels). Hence, no strong arguments can be made about the benefit of one algorithm over the other, which may explain why no single, optimal solution for commercial digital hearing-aids has so far evolved. However, from the current study would appear to suggest that in the case of low input levels, a slow-acting compression should be used with a high compression ratio (i.e., automatic volume control) to provide audibility at this specific input level range, whereas syllabic compression (small compression ratio) or even linear amplification seems to be beneficial at medium to high input levels (comparable results are also found by *Maré* et al. (1992). The case for reducing the compression ratio while increasing the input level is supported by the fact that sensorineural hearing-impairment is combined with a loss of compressive nonlinearity (outer haircell damage) on the basilar membrane, which compresses weak sounds in particular. In addition, the I/O characteristics derived from loudness models for stationary sounds produce the same effect. However, compression limiting should be provided in any case so as to prevent high-level signal peaks.

- »Individual« differences across subjects (such as different residual dynamic ranges and other audiological features of the hearing loss) as well as the practical limitations of the hearing-aid being tested (e.g., limitation in the maximum gain, frequency response and fidelity of the transducer) may play a more important role in some cases than the exact choice of the dynamic compression algorithm. Since the effect of such additional factors was eliminated to a certain degree in the current study (i.e., the same hardware was employed for the same subjects and a very similar fitting rationale was used), certain effects of the dynamic compression algorithm were observable. However, as soon as the additional parameters are imperfectly balanced, (i.e., if two hearing-aids produced by different manufacturers are compared with each other), the effect of the algorithms implemented in the respective hearing-aids can no longer be studied in isolation from the other factors. This clearly limits the validity of comparative field tests that try to compare the user benefit obtained from different hearing-aid hardware and fitting rationales.

Acknowledgements

We would like to thank *Kerstin Sommer* for performing the measurements as well as the subjects who were willing to participate in this study. We would also like to thank two anonymous reviewers and the editor, *Martin Kinkel*, for helpful comments on earlier versions of this manuscript.

This work was supported by the *Bundesministerium für Bildung und Forschung (German Federal Ministry for Research and Education)* and by the *Deutsche Forschungsgemeinschaft (German Research Association).*

References

Appell JE, Gabriel B, Kollmeier B, Marzinzik M, Raß U, Steeger GH, Wittkop T (1999) Implementation and field-test evaluation of simplified noise suppression and dynamic compression algorithms on a dsp prototype hearing aid. Unpublished progress report, Carl von Ossietzky Universität Oldenburg, AG Medizinische Physik

Appell JE, Hohmann V, Kollmeier B (1995) Vergleich verschiedener digital realisierter Signalverarbeitungs-Strategien für Dreikanal-Hörgeräte mit Dynamikkompression. Audiologische Akustik 34 (3), 134–143

Benson D, Clark TM, Johnson JS (1992) Patient experience with multiband full dynamic range compression. Ear and Hear 13 (5), 320–330

Brand T, Kollmeier B (1996) Adaptive Testverfahren in der Audiologie. In: Fortschritte der Akustik – DAGA 96. DEGA e.V., Oldenburg, 60–63

Bustamente PK, Braida LD (1987) Principal-component amplitude compression for hearing impaired. J Acoust Soc Am 82 (4)

Byrne D (1986) The national acoustics laboraties' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. Ear and Hear 7, 257–265

Festen JM (1999) Spectro-temporal modulation transfer and speech intelligibility for multi-band amplitude compression. Handout

Fröhlich T (1993) Digitale Signalverarbeitung für Hörgeräte: Mehrkanalige Lautheitskorrektur im Frequenzbereich. Ph.D. thesis, ETH Zürich

Gatehouse S (1992) The time course and magnitude of perceptual acclimatization to frequency response: Evidence from monaural fitting of hearing aids. J Acoust Soc Am 92, 1258–1268

Genaro Sd, Braida LD, Durlach NI (1986) Multichannel compression for several listeners. J Reh Res Dev 23/1

Hansen M (2000) Einfluss der Kompressionszeitkonstanten auf subjective Sprachverständlichkeit und Klangqualität von Hörgeräten. In Fortschritte der Akustik – DAGA 2000. DEGA e.V., Oldenburg. In Druck

Hohmann V, Kollmeier B (1995a) The effect of multichannel dynamic compression on speech intelligibility. J Acoust Soc Am 97 (2), 1191–1195

Hohmann V, Kollmeier B (1995b) Weiterentwicklung und klinischer Einsatz der Hörfeldskalierung. Audiologische Akustik 34 (2), 48–64

Kießling J, Steffens T (1991) Clinical evaluation of a programmable three channel automatic gain control amplification system. Audiology 30, 70–81

Kollmeier B (1997a) Aktuelle und zukünftige Entwicklungen. In: *Kießling J, Kollmeier B, Diller G* (eds.) Versorgung und Rehabilitation mit Hörgeräten. Georg Thieme Verlag, Stuttgart, 111–131

Kollmeier B (1997b) Hörflächenskalierung – Grundlagen und Anwendung der kategorialen Lautheitsskalierung für Hördiagnostik und Hörgeräteversorgung, vol. 2. Median-Verlag, Heidelberg

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 (4), 2412–2421

Lippmann RP, Braida LD, Durlach NI (1981) Study of multichannel amplitude compression and linear amplification for persons with sensorineural hearing loss. J Acoust Soc Am 69, 524–534

Lunner T, Hellgren J, Arlinger S, Elberling C (1998) Non-linear signal processing in digital hearing aids. Scan Audiol 27 (Suppl. 49), 40–49

Maré MJ, Dreschler WA, Verschuure H (1992) The effects of input-output configuration in syllabic compression on speech perception. J Speech Hear Res 35, 675–685

Marzinzik M, Hohmann V, Appell JE, Kollmeier B (1997) Evaluation of different multi-channel dynamic compression algorithms with regard to recruitment compensation, quality and speech intelligibility. In: *Schick A, Klatte M* (eds.) Seventh Oldenburg Symposium on Psychological Acoustics, vol. 7. BIS, Universität Oldenburg, 619–630

Moore BCJ, Peters RW, Stone MA (1999) Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips. J Acoust Soc Am 105 (1), 400–411

Nábelek JV (1983) Performance of hearing-impaired listeners under various types of amplitude compression. J Acoust Soc Am 74 (3), 776

Neuman A, Bakke M, Mackersie C, Hellman S, Levitt H (1995) Effect of release time in compression hearing aids: Paired comparison judgement of quality. J Acoust Soc Am 98 (6), 3182–3187

Plomp R (1988) The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation transfer function. J Acoust Soc Am 83, 2322–2327

Raß U, Steeger G (2000) A high performance pocket-size system for evaluations in acoustic signal processing. Acustica 86 (2), 374–375

Steinberg JC, Gardner MB (1937) The dependency of hearing impairment on sound intensity. J Acoust Soc Am 9, 11–23

Stone MA, Moore BCJ, Alcantara JI, Glasberg BR (1999) Comparison of different forms of compression using wearable digital hearing aids. J Acoust Soc Am 106 (6), 3603–3619

Stone MA, Moore BCJ, Wojtczak M, Gudgin E (1997) Effects of fast-acting high-frequency compression on the intelligibility of speech in steady and fluctuating background sounds. Brit J Audiol 31 (4), 257–273

Tejero-Calado JC, Rutledge JC, Nelson PB (1998) Combination compression and linear gain processing for digital hearing aids. 20th Annual Int'l. Conf. of the IEEE Engineering in Medicine and Biology Society

Verschuure J, Benning FJ, van Cappellen M, Dreschler WA, Boeremans PP (1998) Speech intelligibility in noise with fast compression hearing aids. Audiology 29 (3), 127–150

Verschuure J, Dreschler WA (1996) Dynamic compression hearing aids. In: *Kollmeier B* (ed.) Psychoacoustics, speech and hearing aids. World Scientific, Singapore, 153–164

Villchur E (1989) Comments on »the negative effect of amplitude compression in multichannel hearing aids in the light of the modulation transfer function«. J Acoust Soc Am 86, 425–427

Walker G, Byrne D (1984) The effect of multiband compression/expansion amplification on the intelligibility of nonsense syllables in noise. J Acoust Soc Am 76 (3), 746

Wesselkamp M, Kliem K, Kollmeier B (1992) Erstellung eines optimierten Satztests in deutscher Sprache. In: *Kollmeier B* (ed.) Moderne Verfahren der Sprachaudiometrie. Buchreihe Audiologische Akustik. Median-Verlag, Heidelberg, 330–343

White MW (1986) Compression systems for hearing aids and cochlear prosthesis. J Reh Res Dev 23/1, 25–39

Working Group on Communication Aids for the Hearing-Impaired (1991) Speech-perception aids for hearing-impaired people: Current status and needed research. J Acoust Soc Am 90, 637–685

6. Oktober 2002 (Freudenstadt)
Homöopathie in der HNO-Heilkunde. Information: Zentralverband der Ärzte für Naturheilverfahren, Am Promenadenplatz 1, D-72250 Freudenstadt. Tel. +49 7441 918580, Fax +49 7441 9185822.

9. Oktober 2002 (Halle)
Multidisziplinäre wissenschaftliche Diskussionssitzung über »Automatisierte Auswertung von Messungen in der Audiologie: von der Signalanalyse zur Befunderstelung« von 10 bis ca.18 Uhr im Universitätsklinikum Halle (Hörsaal der HNO-Klinik). Veranstalter: die Fachgruppe 4.2.1/9.3.2 »Audiologische Akustik« der Informationstechnischen Gesellschaft (ITG) im VDE in Zusammenarbeit mit »Arbeitskreis Audiologie« der Deutschen Gesellschaft für Medizinische Physik (DGMP) und dem Fachausschuss »Hörakustik« der Deutschen Gesellschaft für Akustik (DEGA) sowie der Deutschen gesellschaft für Audiologie (DGA). Weitere Informationen bei *Dr.-Ing. Wolfgang H. Döring*, HNO-Klinik des Universitätsklinikums Aachen, Audiologie, Pauwelsstraße 30, D-52074 Aachen, Tel. +49 (0)241 8088 950/951, Fax +49 (0)241) 8082419, E-Mail: wdoering@ukaachen.de sowie bei *Prof. Dr.-Ing. Thomas Janssen*, Universitäts-HNO-Klinik rechts der Isar, Ismaningerstr. 22, D-81675 München, Tel. +49 (0) 89 41404197, E-Mail: t.janssen@lrz.tu-muenchen.de.

9.–11. Oktober 2002 (Erlangen)
Operations-Fortbildungskurs »Plastischrekonstruktive und ästhetische Nasen- und Ohrmuschelchirurgie«. Wissenschaftliche Leitung: Prof. *Dr. H. Iro, OA Dr. J. Constantinidis*. Information: *OA Dr. J. Constantinidis*, Univ.-HNO-Klinik, Waldstr. 1, D-1054 Erlangen. Tel. +49 9131 853 3156, Fax +49 9131 853 3833.

10.–12. Oktober 2002 (Leipzig)
47. Internationaler Hörgeräte-Akustiker-Kongress. Information: Union der Hörgeräte-Akustiker (UHA), Neubrunnenstr. 3, D-55116 Mainz. Tel. +49 6131 28300, Fax +49 6131 283030 (E-mail: info@uha.de).

1.–2. November 2002 (Erfurt)
Ultraschalldiagnostik im Kopf- Hals-Bereich (A- und B-Bild-Verfahren) sowie Doppler- und Farbdopplersonografie im Hals-Kopf-Bereich. Abschlusskurs einschließlich Dopplersonografie. Information: *Frau R. Hölzer*, Chefsekretarlat, HELIOS Klinikum Erfurt GmbH, Klinik für Hals-, Nasen- und Ohrenheilkunde, Postfach 101263, D-99112 Erfurt. Tel. +49 361 7812101, Fax +49 361 7812102.

8.–9. November 2002 (Dresden)
Allergie-Aufbaukurs. Information: *Dr. B. Hauswald*. Anmeldung: Kongresssekretariat, *Jana Gursinsky*, HNO-Univ.-Klinik Dresden, Fetscherstr. 74, D-01307 Dresden. Tel. +49 351 458 2221, Fax +49 351 458 4326 (E-mail: jana.gursinsky@mailbox.tu-dresden.de).

8.–9. November 2002 (Ulm)
Allergologie als Querschnittsfach. Aufbaukurs Allergologie in der HNO-Heilkunde. Information: Sekretariat *Prof Dr. G. Rettinger*, HNO-Univ.-Klinik, Prittwitzstr. 43, D-89075 Ulm. Tel. +49 731 500 33001, Fax +49 731 500 26703 (E-Mail: ent.department@medizin.uni-ulm.de).

15.–16. November 2002 (Dresden)
Ultraschall-Abschlusskurs A- und B-Mode im Kopf-Hals-Bereich. Information: *Dr. Ch. Offergeld*. Anmeldung: Kongresssekretariat, *Jana Gursinsky*, HNO-Univ.-Klinik Dresden, Fetscherstr. 74, D-01307 Dresden. Tel. +49 351 458 2221, Fax +49 351 458 4326 (E-mail: jana.gursinsky@mailbox.tu-dresden.de).

17. November 2002 (Dresden)
Kurs für Doppier- und Duplexsonografie, Abschlusskurs. Information: *Dr. Ch. Offergeld*. Anmeldung: Kongresssekretariat, *Jana Gursinsky*, HNO-Univ.-Klinik Dresden, Fetscherstr. 74, D-01307 Dresden. Tel. +49 351 458 2221, Fax +49 351 458 4326, E-mail: jana.gursinsky@mailbox.tu-dresden.de

22./23. November 2002 (Köln)
AGERA (Tagung der Arbeitsgemeinschaft ERA) in der Universitäts-HNO-Klinik Köln (Joseph-Stelzmann-Str. 9, 50924 Köln). Hauptthemen : Auditorische Neuropathie, Automatisierte Auswertung, Objektive Diagnostik bei ZAVWS (zentralauditorische Verarbeitungs- und Wahrnehmungsstörungen), Praxisseminar »Diagnostische Nutzung der DPOAE«. Informationen im Internet: www.medizin.uni-koeln.de/kliniken/hno/

28.–30. November (Dresden)
Kurs für Phoniatrie und Pädaudiogie. Themen: Phoniatrie und Pädaudiologie für Weiterbildungsassistenten der HNO. Information: *PD Dr. R. Müller, Dr. E. Müller-Aschoff*. Anmeldung: Kongresssekretariat, *Jana Gursinsky*, HNO-Univ.-Klinik Dresden, Fetscherstr. 74, D-01307 Dresden. Tel. +49 351 458 2221, Fax +49 351 458 4326, E-mail: jana.gursinsky@mailbox.tu-dresden.de

14. Dezember 2002 (Dresden)
8. Interdisziplinäres Allergiesymposium. Information: *Dr. B. Hauswald*. Anmeldung: Kongresssekretariat, *Jana Gursinsky*, HNO-Univ.-Klinik Dresden, Fetscherstr. 74, D-01307 Dresden. Tel. +49 351 458 2221, Fax +49 351 458 4326, E-mail: jana.gursinsky@mailbox.tu-dresden.de

26.Februar–1. März 2003 (Hannover)
Otology Update Kurs. 6. Operationskurs ür Mittelohr- und Schädelbasischirurgie inkl. Mittel- und Innenohrimplantate mit Live-OPs und praktischen Übungen. Information: Gabi Richardson, HNO-Klinik der Medizinischen Hochschule Hannover, Carl-Neuberg-Str. 1, D-30625 Hannover, Phone +49 (0)511 5239161, Fax +49 (0)511 5233293, E-Mail: ric@hno.mh-hannover.de

VITAE AUTORUM

Mark Marzinzik, Dipl.-Phys, Dr. rer. nat.; geboren 1970 in Bremen. Von 1990 bis 1996 Studium der Physik an der Carl von Ossietzky Universität Oldenburg. 2000 Promotion bei Prof. Dr. Dr. Birger Kollmeier zum Thema »Noise reduction schemes for digital hearing aids and their use for the hearing impaired«. Von 1996 bis 2001 wiss. Mitarbeiter in der Arbeitsgruppe Medizinische Physik. E-Mail: mark@medi.physik.uni-oldenburg.de

Mark Marzinzik, Dipl.-Phys, Dr. rer. nat.; born 1970 in Bremen. Studied physics from 1990 to 1996 at the Carl von Ossietzky University Oldenburg. Ph.D. degree in 2000 (supervised by Prof. Dr. Dr. Birger Kollmeier) with a dissertation on »Noise reduction schemes for digital hearing aids and their use for the hearing impaired«. Worked from 1996 to 2001 as a Research Associate with the Medical Physics Department, Universität Oldenburg (E-Mail: mark@medi.physik.uni-oldenburg.de).

Jens-Ekkehart Appell, Diplom-Physiker, Dr. rer. nat.: Zur Vita autoris/for biographical notes see »Zeitschrift für Audiologie/Audiological Acoustics« 4/2001, Anschrift/address: OFFIS – F&E Bereich Eingebettete Systeme, Escherweg 2, D-26121 Oldenburg (E-mail: jens-e.appell@offis.de)

Birgitta Gabriel, Dr. rer. nat.: Zur Vita autoris/for biographical notes see »Zeitschrift für Audiologie/Audiological Acoustics« 3/2000, Anschrift/address: Carl von Ossietzky Universität, AG Medizinische Physik, FB8/Physik, Carl-von-Ossietzky-Str. 9–11, D-26111 Oldenburg (E-mail: B.Gabriel@Hoerzentrum-Oldenburg.de).

Volker Hohmann, Dr. rer. nat.: Zur Vita autoris/for biographical notes see »Zeitschrift für Audiologie/Audiological Acoustics« 4/2001, Anschrift/address: Carl von Ossietzky Universität, AG Medizinische Physik, FB8/Physik, Carl-von-Ossietzky-Str. 9–11, D-26111 Oldenburg (E-mail: vh@medi.physik.uni-oldenburg.de).

Birger Kollmeier, Prof. Dr. Dr.: Zur Vita autoris/for biographical notes see »Zeitschrift für Audiologie/Audiological Acoustics« 4/2001, Anschrift/address: Carl von Ossietzky Universität, AG Medizinische Physik, FB8/Physik, Carl-von-Ossietzky-Str. 9–11, D-26111 Oldenburg (E-mail: birger.kollmeier@uni-oldenburg.de).

Inserenten dieses Heftes

Median-Verlag
Phonak
Widex

STELLENMARKT

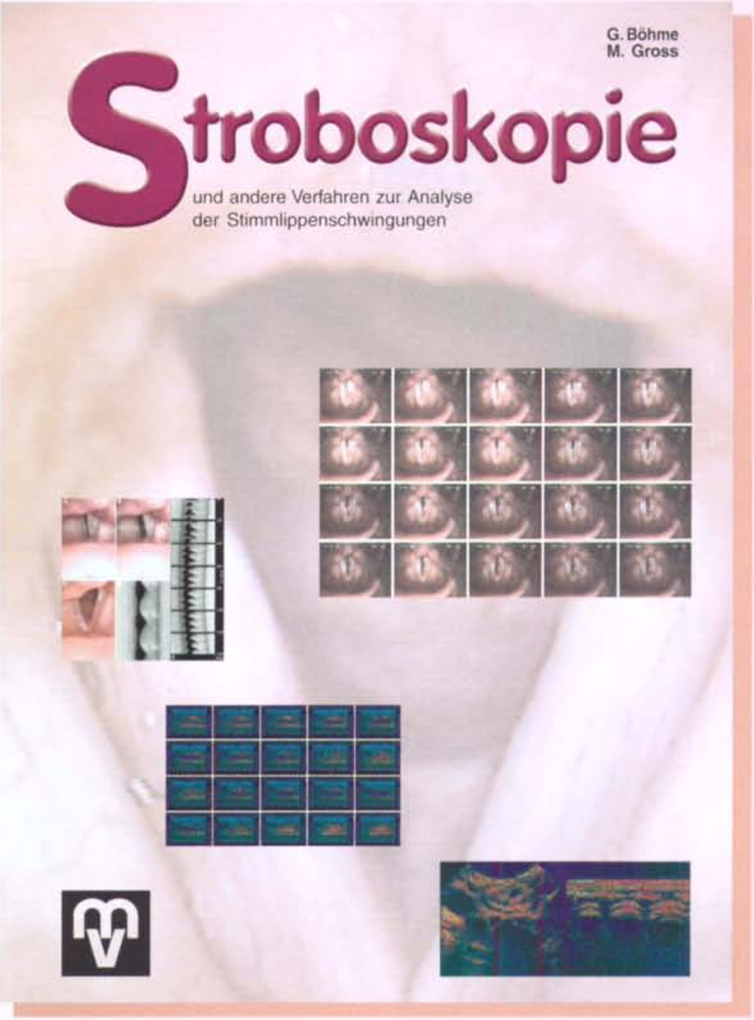
Stroboskopie und andere Verfahren zur Analyse der Stimmlippenschwingungen

von Gerhard Böhme und Manfred Gross

2001, 200 Seiten, 126 teilweise vierfarbige Abbildungen,
Hardcover, ISBN 3-922766-69-2 · € 51,– / sfr 90,–

Aus dem Vorwort:

Wir stellten uns deshalb die Aufgabe, nach Schilderung der theoretischen Grundlagen den klinischen Wert bei funktionellen und organischen Erkrankungen des Kehlkopfes zu beschreiben. Die Laryngo-Stroboskopie ist eine Routine-Untersuchung. Es ist ein intensives Training unter Anleitung notwendig, bis eine hohe Sicherheit in der stroboskopischen Differentialdiagnostik erreicht wird.



Aus heutiger Sicht kann die Laryngo-Stroboskopie keineswegs alle Fragen zur Analyse der Stimmlippenschwingungen beantworten. Deshalb gewinnen weitere Verfahren zur Schwingungsanalyse der Glottis zunehmend an Bedeutung. An erster Stelle muss die Hochgeschwindigkeitskinematographie genannt werden. Damit gelangt man zu einer präzisen Darstellung des gesamten Stimmlippenschwingungsablaufes, unabhängig von Aperiodizitäten, Irregularitäten, Phasenverschiebungen etc. Aber auch andere Methoden, besonders die Videokymographie und Elektroglottographie eignen sich zur Schwingungsanalyse der Glottis. Dagegen erfordern die Ultraschalldiagnostik der Glottis und ihre Beziehungen zu den Schwingungsaktivitäten der Stimmlippen noch weitere intensive Forschungsarbeit.

Ein Kernstück des Lehrbuches sind unzweifelhaft die videostroboskopischen Sequenzen, bei denen die Einzelbilder je einer charakteristischen Schwingungsperiode nebeneinander dargestellt sind. Das visuelle Wahrnehmungssystem des Menschen besitzt außerordentlich stark ausgebildete Fähigkeiten, auch winzige Veränderungen in einem dynamischen, bewegten Ablauf zu erkennen, während der Vergleich statischer Bilder wesentlich schwerer fällt.

Bestellen Sie beim:
Median-Verlag von Killisch-Horn GmbH
Postfach 10 39 64 · 69029 Heidelberg
Tel. 0 62 21/90 50 915 · Fax 0 62 21/90 50 920
E-mail: median-verlag@t-online.de

Auslieferung Schweiz:
Roggen-Amrein
AG für Kommunikation und Marketing
Rainstraße 2a · CH-5415 Nussbaumen
Tel. 056/282 25 65 · Fax 056/282 25 33