

# Zur Lautheit fließender Sprache

Hugo Fastl\*

*Zusammenfassung Für den Zusammenhang zwischen der subjektiv wahrgenommenen Lautstärke (Lautheit) fließender Sprache und den objektiv meßbaren Schallparametern, Schallpegel bzw. Bandbreite werden einfache Richtwerte angegeben. Eine Halbierung der Lautheit fließender Sprache ergibt sich bei einer Pegelerniedrigung um etwa 7 dB, bei einer Tiefpaßbegrenzung mit etwa 700 Hz Grenzfrequenz oder bei einer Hochpaßbegrenzung mit etwa 1000 Hz Grenzfrequenz. Durch Amplitudenbegrenzung verzerrte fließende Sprache und unverzerrte Sprache werden bei gleichem maximalem Schallpegel als nahezu gleich laut wahrgenommen. Sprachsimulierendes Rauschen und fließende Sprache erzeugen dann die gleiche Lautheit, wenn der Schallpegel des Rauschens dem Maximalwert des stark zeitvarianten Schallpegels der Sprache entspricht.*

## 1. Einleitung

Die Lautstärkeempfindung (Lautheit), die ein Schall hervorruft, wird vor allem von seinem Schallpegel bestimmt. Darüber hinaus sind jedoch die Spektralverteilung sowie die Zeitstruktur von Schallen für deren Lautheit von entscheidender Bedeutung (vgl. [1], [2]). Bei fließender Sprache handelt es sich um einen sehr komplexen Schall mit ausgeprägten zeitlichen Schwankungen des Schallpegels und stark zeitvarianten Spektralverteilungen. Im Hinblick auf die Lautheit fließender Sprache interessieren insbesondere drei Fragestellungen:

- 1) Auf welche Weise beeinflussen die Schallparameter Pegel und Bandbreite die Lautstärkeempfindung für fließende Sprache?
- 2) Wie wirken sich bei Amplitudenbegrenzung auftretende Verzerrungen auf die Lautheit fließender Sprache aus?
- 3) Läßt sich die Lautheit fließender Sprache anhand einer objektiven Meßgröße (zum Beispiel Schalldruckpegel) beschreiben?

Eine allgemein gültige, umfassende Antwort auf diese Fragen konnte trotz umfangreicher, systematischer Untersuchungen in zahlreichen Forschungslaboratorien bisher nicht gefunden werden (für Literaturhinweise vgl. [3], [4], [5], [6]). Um dennoch wenigstens praxisorientierte Anhaltswerte angeben zu können, wurden die im folgenden beschriebenen Meßreihen durchgeführt. Dabei wurden zugunsten von einfachen, für den Anwender als Maßstab dienenden Richtwerten bewußt Detailprobleme außer acht gelassen. Mittels psychoakustischer Experimente wurde die Abnahme der Lautheit eines Test-Satzes bei Pegelerniedrigung, sowie bei Tiefpaß- bzw. Hochpaßbegrenzung bestimmt.

Ferner wurde die Lautheit eines durch Amplitudenbegrenzung verzerrten Test-Satzes relativ zur Lautheit des unverzerrten Satzes beurteilt. Schließlich wurde die Lautheit des

## 1. Introduction

The auditory impression of the strength of sound (loudness) is mainly determined by the sound level. In addition, the spectral distribution as well as the temporal structure of sounds are of decisive importance for their loudness (see [1], [2]). Running speech is made up of very complex sound with pronounced temporal fluctuations of the sound level and spectral distributions which strongly vary with time. As to the loudness of running speech it is mainly three questions that are of interest:

- 1) In what manner is the loudness perception for running speech influenced by the parameters sound level and bandwidth?
- 2) What are the effects of distortions caused by peak clipping on the loudness of running speech?
- 3) Can the loudness of running speech be described by an objectively measurable parameter (for instance, the sound-pressure level)?

Despite extensive and systematic investigations in numerous research laboratories it has not been possible to find a generally valid and comprehensive answer to these questions (see bibliography [3], [4], [5], [6]). However, in order to furnish at least some practice-oriented guide values, the test series described later in the text were carried out. In doing this, secondary problems were left out of account in order to arrive at simple standard values which may serve as a yardstick to the user. By means of psycho-acoustic experiments the decrease in loudness of a test sentence was determined at reduced levels and with low-pass and high-pass filtering. Moreover, the loudness of a test sentence, distorted by peak clipping was judged relative to the loudness of the undistorted sentence. Finally, the loudness of a test sentence was compared with the loudness of speech-simulating broadband noise.

\* Lehrstuhl für Elektroakustik der Technischen Universität München

\* Institute for Electroacoustics, Technical University Munich

# Loudness of running Speech

**Summary** Simple guide values are indicated for the relation between the subjectively perceived volume (loudness) of running speech and such objectively measurable sound parameters as sound level and bandwidth. The loudness of running speech is halved when the level is reduced by about 7 dB or when low-pass filtering with a cut-off frequency of roughly 700 Hz or high-pass filtering with a cut-off frequency of about 1,000 Hz is provided. Running speech, distorted by peak clipping and undistorted speech of equal maximal sound level are perceived as being nearly equally loud. Speech-simulating noise and running speech produce the same loudness whenever the noise level corresponds to the maximum value of the strongly time-dependent level of speech.

Test-Sätze mit der Lautheit eines sprachsimulierenden Breitbandrauschens verglichen.

## 2. Experimente

Um praxisbezogene Richtwerte zu erhalten, wurden sämtliche Lautheitsvergleiche weder im echoarmen Raum noch bei Kopfhörerdarbietung durchgeführt; vielmehr wurden die Schalle in der Bibliothek des Lehrstuhls für Elektroakustik der Technischen Universität München über eine handelsübliche Hi-Fi-Anlage wiedergegeben. Der Bibliotheksraum weist eine Grundfläche von etwa 11 m × 4,50 m (Volumen etwa 140 m<sup>3</sup>) auf, die Nachhallzeit bei mittleren Frequenzen liegt um 600 ms, der Störschallpegel beträgt im Mittel 40 dB(A). Als ein jederzeit auch an anderer Stelle reproduzierbares Beispiel für fließende Sprache wurde folgender Test-Satz aus einem Bezugsband (Typ 4730-72) der Agfa-Gevaert AG gewählt:

»Bezugsband 19s, nach DIN 45513 Blatt 3«.

Für die Versuchsserien wurden jeweils 24 Paare dieses Satzes mit einem Studio-Tonbandgerät aufgenommen. Während der erste Satz eines Paares stets unverändert blieb, wurde der jeweils zweite Satz je nach Versuchsreihe entweder mit anderem Pegel dargeboten oder tiefpaßbegrenzt bzw. hochpaßbegrenzt. Zwischen beiden Sätzen (gleichen Textes) lag eine Pause von etwa 700 ms; die Pause zwischen den Satzpaaren war etwa 3 s lang. Die Darbietungslautstärke wurde so gewählt, daß sich in der Bibliothek für den jeweils ersten Satz eines Paares ein maximaler Schallpegel von etwa 70 dB (A-Bewertung, »fast«) ergab.

Zwischen 8 und 16 normalhörende Versuchspersonen im Alter von 25 bis 52 Jahren nahmen an den Experimenten teil. Die Gruppe umfaßte sowohl Personen mit jahrelanger Übung in psychoakustischen Experimenten, wie auch völlig ungeübte Beobachter. Bei den Lautheitsvergleichen hatten die Versuchspersonen folgendermaßen vorzugehen:

Der subjektiv wahrgenommenen Lautstärke (Lautheit) des jeweils ersten Satzes eines Paares wurde der Zahlenwert  $N_{ref} = 100$  zugeordnet. Relativ zu diesem Zahlenwert war die Lautheit des jeweils zweiten Satzes eines Paares zu skalieren.

## 2. Experiments

In order to obtain practice-oriented guide values, all loudness comparisons were carried out neither in the anechoic chamber nor was speech presented via headphones; on the contrary, sound was reproduced by a commercially available hi-fi system in the library of the Institute for Electroacoustics of the Technical University of Munich. The library has an area of about 11 m × 4.50 m (a volume of roughly 140 m<sup>3</sup>), the reverberation time at mean frequencies is about 600 ms, the noise level averages 40 dB(A). As an example of running speech which can be reproduced at any time and place, the following test sentence (in German) was selected from a reference tape (type 4730-72) of Agfa-Gevaert AG: "Bezugsband 19 s, nach DIN 45513 Blatt 3" (Reference tape 19 s, according to German Industrial Norm 45513, sheet 3).

For the various test series 24 pairs of this sentence were recorded with a studio tape recorder. While the first sentence of a pair remained unchanged, the second sentence was, depending on the test series, either presented at different levels or low-pass filtered or high-pass filtered. Between the two sentences (of the same text) there was an interval of about 700 ms; the interval between the sentence pairs was roughly 3 s. The volume, at which the sentences were presented in the library was selected in such a manner that a maximum sound level of about 70 dB (A-weighting, "fast") was obtained for the respective first sentence of a pair.

Between 8 to 16 normally hearing test persons aged from 25 to 52 years were chosen to participate in the experiments. The group comprised both persons with years of experience in psycho-acoustic tests as well as completely untrained participants. For the loudness comparison the test persons had to proceed in the following manner:

The numerical value  $N_{ref} = 100$  was correlated to the subjectively perceived volume (loudness) of the first sentence of a pair. The loudness of the second sentence of a pair had to be scaled in relation to this value. If, for instance, the second sentence was subjectively perceived as being only half as loud

Wurde beispielsweise der zweite Satz subjektiv als halb so laut empfunden wie der erste Satz, so sollte der Zahlenwert  $N = 50$  angegeben werden, bei subjektiv gleich lauten Sätzen der Zahlenwert  $N = 100$ , bei Abnahme der Lautheit des zweiten Satzes auf  $\frac{1}{3}$  der Lautheit des ersten Satzes der Zahlenwert  $N = 20$  usw. Nach jedem Satzpaar hatte jede Versuchsperson den Zahlenwert, der die Relation der Lautheiten beider Sätze beschreibt, in eine Antwortkarte einzutragen.

## 2.1 Pegelerniedrigung

In diesem Experiment wurde der Schallpegel des jeweils zweiten Satzes eines Paares gegenüber dem Pegel des ersten Satzes um eine Pegeldifferenz  $\Delta L = -2, -4, -6, -8, -10$  oder  $-12$  dB abgesenkt. Die Aufeinanderfolge der verschiedenen Pegelabsenkungen war nicht nach Pegelstufen geordnet, sondern zufällig. Acht Versuchspersonen waren an dieser Messung beteiligt, insgesamt wurden 192 Lautheitsvergleiche durchgeführt. Die Ergebnisse zeigt Abb. 1; die relative Lautheit  $N/N_{ref}$  des Satzes mit geringerem Pegel ist als Funktion der Pegelerniedrigung  $\Delta L$  dargestellt. Aus jeweils 32 Meßpunkten wurden die Zentralwerte mit wahrscheinlicher Schwankung gebildet, d. h. innerhalb des durch vertikale Striche angedeuteten Bereiches liegen 50 Prozent aller Aussagen der Versuchspersonen (vgl. [7]).

Die in Abb. 1 dargestellten Ergebnisse weisen darauf hin, daß für die Halbierung der Lautheit fließender Sprache eine Schallpegelabsenkung um etwa 7 dB notwendig ist. Bei einer Pegelabsenkung um 10 dB verringert sich die subjektiv wahrgenommene Lautstärke (Lautheit) fließender Sprache auf etwa  $\frac{1}{3}$ . Die geringen wahrscheinlichen Schwankungen in

as the first sentence, the test person had to indicate the value  $N = 50$ , when sentences were subjectively perceived as being of equal loudness the value  $N = 100$  and, when the loudness of the second sentence dropped to  $\frac{1}{3}$  of the loudness of the first sentence the value  $N = 20$  and so on. After each pair of sentences the test person had to enter the value of the loudness relationship between the two sentences in a reply card.

## 2.1 Decrease in level

In this experiment the sound level of the second sentence of a pair was decreased by a level difference of  $\Delta L = -2, -4, -6, -8, -10$  or  $-12$  dB compared with the level of the first sentence. The sequence of the various level decreases were not arranged in steps but left to chance. Eight test persons were examined and a total of 192 loudness comparisons were made. The results are indicated in Fig. 1; the relative loudness  $N/N_{ref}$  of the sentence with the lower level is represented as a function of the level decrease  $\Delta L$ . Medians with interquartile ranges were derived from 32 data points, respectively, i. e. 50 per cent of all statements made by the test persons fall within the range marked by vertical strokes (cf. [7]).

The results shown in Fig. 1 indicate that the sound level must be decreased by about 7 dB to halve the loudness of running speech. With a level decrease of 10 dB, the subjectively perceived volume (loudness) of running speech drops to about  $\frac{1}{3}$ . The small interquartile ranges in Fig. 1 suggest that for all test persons, trained or untrained, the loudness of running speech diminishes roughly to the same extent with a given decrease in the sound level. An evaluation of the statements made by each individual test person shows that

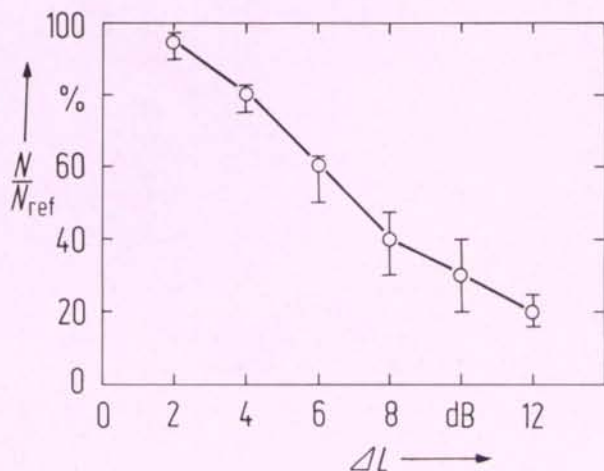


Abb. 1 Abnahme der relativen Lautheit  $N/N_{ref}$  von fließender Sprache bei Erniedrigung des Schallpegels um  $\Delta L$ .

Fig. 1 Decrease of the relative loudness  $N/N_{ref}$  of running speech when the sound level is lowered by  $\Delta L$ .

Abb. 1 deuten darauf hin, daß für alle Versuchspersonen – ob geübt oder ungeübt – bei einer vorgegebenen Schallpegelabsenkung die Lautheit fließender Sprache etwa in gleichem Maße abnimmt. Die Auswertung der Angaben jeder einzelnen Versuchsperson ergibt, daß zwischen den vier je Pegeldifferenz  $\Delta L$  angegebenen Zahlenwerten  $N$  größere Differenzen als 10 nur selten vorkommen. Das bedeutet, daß die Lautheitsvergleiche von den einzelnen Versuchspersonen gut reproduziert werden können und daß hinsichtlich der Abnahme der Lautheit fließender Sprache keine großen individuellen Unterschiede bestehen.

Die angegebene Schallpegeldifferenz von 7 dB zur Halbierung der Lautheit fließender Sprache stimmt sehr gut mit Daten aus der Literatur bei Darbietung in normalen Räumen überein ([4], [6]); bei Darbietung unter Laborbedingungen (zum Beispiel echoarmer Raum) können sich jedoch andere Werte ergeben (vgl. [8], [9]).

## 2.2 Tiefpaßbegrenzung

In dieser Versuchsreihe wurde der Einfluß einer Tiefpaßbegrenzung bei  $f_g = 180, 250, 355, 500, 1000, 2000, 4000$  oder  $8000$  Hz auf die Lautheit fließender Sprache untersucht. Die Tiefpässe wiesen eine Flankensteilheit von 35 dB/Oktave auf. Entsprechend der vorhergehenden Versuchsreihe wurden Paare aus *nicht* tiefpaßbegrenztem Test-Satz und tiefpaßbegrenztem Test-Satz dargeboten. Die Versuchspersonen wurden besonders darauf hingewiesen, lediglich *Lautheitsunterschiede* zu skalieren, nicht etwa aber Sprachverständlichkeit, Hifi-Qualität, Sprachgüte oder dergleichen.

Zwölf Versuchspersonen waren an dieser Meßreihe beteiligt;

among the four values  $N$  indicated for each level difference  $\Delta L$ , differences higher than 10 are rarely found. This means that the loudness comparison tests can be readily reproduced by the individual examinees and that no great individual differences exist with respect to the decrease in loudness of running speech

The sound level difference of 7 dB required to halve the loudness of running speech agrees very well with the data in literature when speech is presented in normal rooms ([4], [6]); however, when speech is presented under laboratory conditions (e.g. anechoic chamber) other values may be obtained (see [8], [9]).

## 2.2 Low-pass filtering

In this test series the influence of low-pass filtering at a cut-off frequency  $f_g = 180, 250, 355, 500, 1000, 2000, 4000$  or  $8000$  Hz on the loudness of running speech was investigated. The low-pass filters had a slope of 35 dB/octave. In accordance with the preceding test series, pairs from the test sentence *without* low-pass filtering and from the low-pass filtered test sentence were presented. The test persons were explicitly instructed to assess only the *differences in loudness* but not speech intelligibility, high-fidelity, quality of speech or other factors.

This series which comprised a total of 288 loudness comparison tests was run on 12 examinees. The results are indicated in Fig. 2 as medians with interquartile ranges each derived from 36 data points. The relative loudness  $N/N_{ref}$  of the low-pass filtered test sentence is shown as function of the low-pass cut-off frequency  $f_g$ .

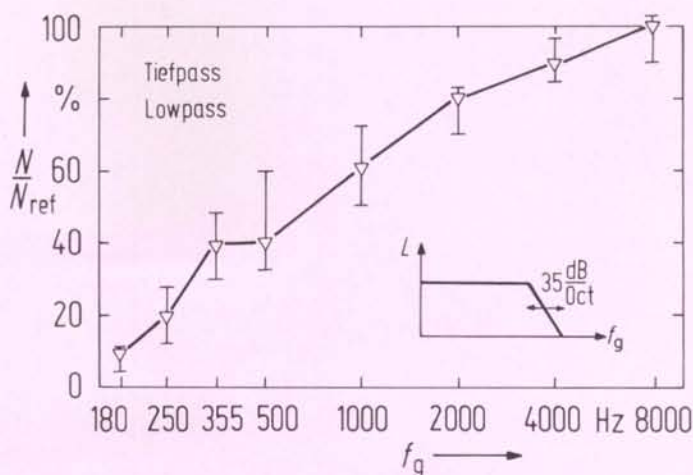


Abb. 2 Abnahme der relativen Lautheit  $N/N_{ref}$  von fließender Sprache für Tiefpaßbegrenzung bei  $f_g$

Fig. 2 Decrease of the relative loudness  $N/N_{ref}$  of running speech for low-pass filtering at  $f_g$

sie führten insgesamt 288 Lautheitsvergleiche durch. Die Ergebnisse sind in Abb. 2 als Zentralwerte mit wahrscheinlichen Schwankungen, abgeleitet aus jeweils 36 Meßdaten, angegeben. Die relative Lautheit  $N/N_{ref}$  des tiefpaßbegrenzten Test-Satzes ist als Funktion der Tiefpaßgrenzfrequenz  $f_g$  dargestellt.

Die in Abb. 2 gezeigten Ergebnisse legen den Schluß nahe, daß zur Halbierung der Lautheit des Test-Satzes eine Bandbegrenzung auf etwa 700 Hz erforderlich ist. Andererseits erniedrigt sich bei einer Bandbegrenzung auf 4000 Hz die Lautheit fließender Sprache kaum. Eine obere Bandgrenze, wie sie im Fernsprechnetz oder bei Mittelwellensendern üblich ist, beeinflußt offensichtlich die subjektiv wahrgenommene Lautstärke (Lautheit) fließender Sprache nur unwesentlich. Dagegen ergeben sich bei Tiefpaßbegrenzung von Sprache deutliche Unterschiede hinsichtlich Hifi-Qualität oder Sprachgüte. Im allgemeinen lieferten die Versuchspersonen gut reproduzierbare Aussagen mit nur geringen individuellen Unterschieden; lediglich bei  $f_g = 500$  Hz unterscheiden sich die individuellen Angaben deutlicher als sonst üblich.

### 2.3 Hochpaßbegrenzung

In dieser Meßreihe wurde der Einfluß einer Hochpaßbegrenzung bei  $f_g = 180, 250, 355, 500, 710, 1000, 2000$  oder  $4000$  Hz auf die Lautheit fließender Sprache untersucht. Die Flankensteilheit der Hochpaßfilter war ebenfalls  $35$  dB/Oktave. Die 12 Versuchspersonen wurden wieder ausdrücklich darauf hingewiesen, beim Vergleich der Paare aus nicht hochpaßbegrenztem Test-Satz und hochpaßbegrenztem Test-Satz, ausschließlich die Lautheit zu bewerten und von

The results shown in Fig. 2 suggest that a limitation of the bandwidth to about 700 Hz is required to halve the loudness of the test sentence. On the other hand, limiting the bandwidth to 4,000 Hz hardly causes the loudness of running speech to decrease. An upper band limit such as that in general use with telephone systems or medium-frequency transmitters, appears to affect the subjectively perceived loudness of running speech only to a negligible extent. By contrast, low-pass filtering of speech produces distinct differences in hi-fi quality or quality of speech. In general the test persons furnish readily reproducible data with only slight variations from person to person; only at a cut-off frequency of 500 Hz do the individual values differ to a greater extent than is normally the case.

### 2.3 High-pass filtering

In this test series the influence of high-pass filtering at  $f_g = 180, 250, 355, 500, 710, 1000, 2000$  or  $4000$  Hz on the loudness of running speech was investigated. The slope of the high-pass filters was also  $35$  dB/octave. The 12 test persons were expressly instructed to assess loudness exclusively on comparing the pairs from the non-high-pass filtered test sentence and from the high-pass filtered test sentence, and to forget all other attributes such as sound pattern, speech intelligibility so forth.

Fig. 3 shows the results of 288 loudness comparison tests. The relative loudness  $N/N_{ref}$  of the high-pass filtered test sentence is displayed as a function of the high-pass cut-off frequency  $f_g$ . The symbols represent medians with interquartile ranges each derived from 36 data points.

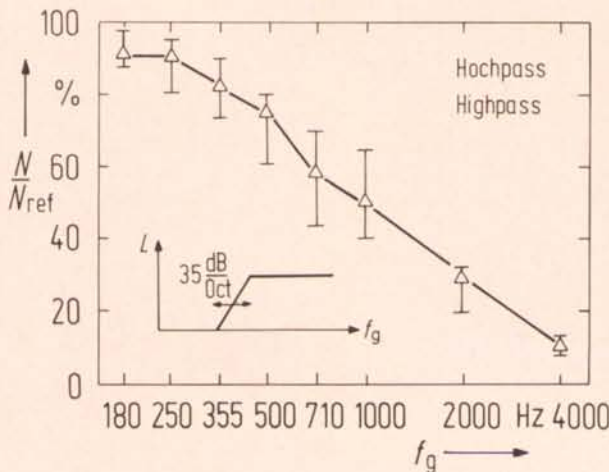


Abb. 3 Abnahme der relativen Lautheit  $N/N_{ref}$  von fließender Sprache für Hochpaßbegrenzung bei  $f_g$

Fig. 3 Decrease of the relative loudness  $N/N_{ref}$  of running speech for high-pass filtering at  $f_g$

allen anderen Attributen wie Klangbild, Sprachverständlichkeit usw. zu abstrahieren.

Abb. 3 zeigt die Ergebnisse von 288 Lautheitsvergleichen. Die relative Lautheit  $N/N_{ref}$  des hochpaßbegrenzten Test-Satzes ist als Funktion der Hochpaßgrenzfrequenz  $f_g$  dargestellt. Die Symbole repräsentieren Zentralwerte mit wahrscheinlichen Schwankungen, abgeleitet aus je 36 Meßpunkten.

Gemäß Abb. 3 ergibt sich eine Halbierung der Lautheit des Test-Satzes bei einer Hochpaßgrenzfrequenz von etwa 1000 Hz. Für die im Fernsprecherkehr bedeutsame untere Grenzfrequenz  $f_g = 355$  Hz zeigt sich eine Lautheitseinbuße von lediglich etwa 20 Prozent. Besonders deutlich wird der Einfluß des angelegten Kriteriums auf die Aussagen der Versuchspersonen für den Meßpunkt bei  $f_g = 4000$  Hz: während die Sprachverständlichkeit des Test-Satzes bei dieser Hochpaßbegrenzung nahezu vollständig erhalten bleibt (der Satz klingt wie geflüstert), nimmt die Lautheit ganz erheblich ab. Ähnlich wie in den oben beschriebenen Experimenten liefern die einzelnen Versuchspersonen gut reproduzierbare Daten; es zeigen sich keine ausgeprägten individuellen Unterschiede.

## 2.4 Amplitudenbegrenzung

In dieser Meßreihe wurde durch Amplitudenbegrenzung verzerrte fließende Sprache hinsichtlich ihrer Lautheit mit unverzerrter Sprache verglichen. Die Amplitudenbegrenzung wurde mit zwei Zener-Dioden durchgeführt; am Ausgang des Begrenzers ergab sich für den übersteuerten, amplitudenbegrenzten Test-Satz – relativ zu einer linearen Übertragung – ein um etwa 15 dB geringerer maximaler Schallpegel. Den 8 Versuchspersonen wurden 24 Paare aus verzerrtem und unverzerrtem Test-Satz dargeboten. Der maximale Schallpegel des verzerrten Test-Satzes war immer auf 70 dB begrenzt (A-Bewertung, „fast“), die Schallpegelmaxima des unverzerrten Test-Satzes nahmen Werte von  $L_{max} = 62, 65, 68, 71, 74$  und  $77$  dB(A) an. Wie in den vorhergehenden Experimenten wurden die Versuchspersonen angewiesen, ausschließlich die Lautheit von verzerrtem und unverzerrtem Test-Satz zu vergleichen.

In Abb. 4 sind die Ergebnisse von 192 Lautheitsvergleichen als Zentralwerte mit wahrscheinlichen Schwankungen, abgeleitet aus jeweils 32 Meßpunkten, angegeben. Die relative Lautheit  $N/N_{ref}$  des *nicht* amplitudenbegrenzten Test-Satzes ist als Funktion seines maximalen Schallpegels  $L_{max}$  dargestellt.

Gemäß Abb. 4 ergibt sich für den durch Amplitudenbegrenzung verzerrten und den unverzerrten Test-Satz die gleiche Lautheit ( $N/N_{ref} = 100$  Prozent) bei einem maximalen Schallpegel  $L_{max} = 71$  dB(A) des unverzerrten Test-Satzes. Da der maximale Schallpegel des amplitudenbegrenzten Test-Satzes bei 70 dB(A) lag, läßt dieses Ergebnis folgende

Angewandt auf Fig. 3 die Lautheit der Test-Satzes ist halbiert bei einer Hochpaßcut-off-Frequenz von etwa 1000 Hz. Bei der niedrigeren Cut-off-Frequenz  $f_g = 355$  Hz, die für die Fernsprechanwendung von Bedeutung ist, beträgt die Lautheitsminderung nur 20 Prozent. Der Einfluss des Kriteriums auf die Aussagen der Versuchspersonen ist besonders eindrucksvoll bei der Messung bei der Cut-off-Frequenz  $f_g = 4000$  Hz: Die Sprachverständlichkeit für den Test-Satz bleibt bei dieser Hochpaßcut-off-Frequenz nahezu unverändert (der Satz klingt wie geflüstert), die Lautheit sinkt jedoch beträchtlich. Wie bei den oben beschriebenen Experimenten lieferten die einzelnen Versuchspersonen gut reproduzierbare Daten; keine markanten Unterschiede zwischen den Personen wurden festgestellt.

## 2.4 Peak clipping

In dieser Testreihe wurde fließende Sprache, die durch Peakclipping verzerrt wurde, hinsichtlich ihrer Lautheit mit unverzerrter Sprache verglichen. Die Amplitudenbegrenzung wurde durch zwei Zener-Dioden realisiert; am Ausgang des Limiters wurde die Peakclipping-Verzerrung erreicht, die den maximalen Schallpegel des Peakclipping-Test-Satzes um etwa 15 dB gegenüber dem unverzerrten Test-Satz senkte. Den 8 Versuchspersonen wurden 24 Paare aus Peakclipping- und unverzerrtem Test-Satz dargeboten. Der maximale Schallpegel des Peakclipping-Test-Satzes war immer auf 70 dB begrenzt (A-Bewertung, „fast“), die Schallpegelmaxima des unverzerrten Test-Satzes nahmen Werte von  $L_{max} = 62, 65, 68, 71, 74$  und  $77$  dB(A) an. Wie in den vorhergehenden Experimenten wurden die Versuchspersonen angewiesen, ausschließlich die Lautheit von Peakclipping- und unverzerrtem Test-Satz zu vergleichen.

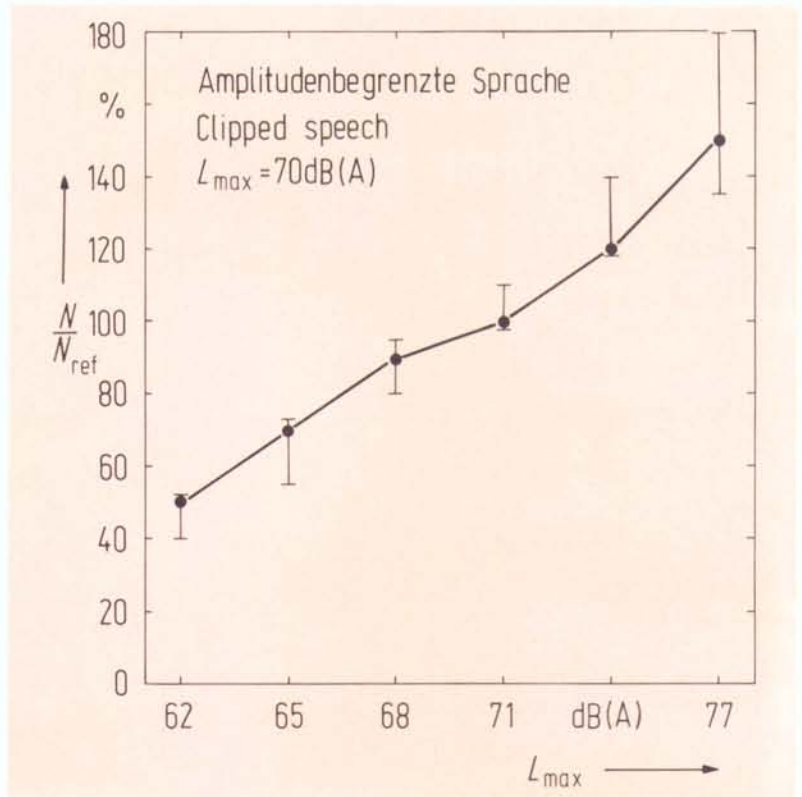
In Fig. 4 sind die Ergebnisse von 192 Lautheitsvergleichen als Zentralwerte mit Interquartilsbereichen, abgeleitet aus jeweils 32 Meßpunkten, angegeben. Die relative Lautheit  $N/N_{ref}$  des *nicht* Peakclipping-Test-Satzes ist als Funktion seines maximalen Schallpegels  $L_{max}$  dargestellt.

Angewandt auf Fig. 4, gleiche Lautheit ( $N/N_{ref} = 100$  Prozent) der Test-Satzes, die durch Peakclipping verzerrt wurde, und der unverzerrten Test-Satzes wird bei einem maximalen Schallpegel  $L_{max} = 71$  dB(A) des unverzerrten Test-Satzes erreicht. Da der maximale Schallpegel des Peakclipping-Test-Satzes bei 70 dB(A) lag, führt dies zu dem folgenden Schluss: Peakclipping-Verzerrung (15 dB) und nicht Peakclipping-Verzerrung erzeugen bei gleichem maximalen Schallpegel (A-Bewertung, „fast“) nahezu die gleiche Lautheit.

Abgesehen von den Werten bei  $L_{max} = 77$  dB(A), lieferten die Versuchspersonen gut reproduzierbare Daten mit nur geringen Schwankungen zwischen den Personen. Im Vergleich mit den früheren Experimenten, bei denen die Lautheitsvergleiche zwischen Peakclipping- und unverzerrter Sprache als schwieriger empfunden wurden, waren die Versuchspersonen nicht sicher, ob sie die Lautheit der Peakclipping- und unverzerrten Sprache immer richtig verglichen hatten.

Abb. 4 Lautheitsvergleich von durch Amplitudenbegrenzung verzerrter mit unverzerrter fließender Sprache. Relative Lautheit  $N/N_{ref}$  der unverzerrten Sprache als Funktion ihres maximalen, A-bewerteten Schallpegels  $L_{max}$

Fig. 4 Loudness comparison of running speech, distorted by peak clipping and undistorted speech. Relative loudness  $N/N_{ref}$  of the undistorted speech as a function of its maximal, A-weighted sound level  $L_{max}$



Schlußfolgerung zu: Signifikant amplitudenbegrenzte fließende Sprache (Übersteuerung 15 dB) und nicht amplitudenbegrenzte Sprache rufen bei gleichem maximalem Schallpegel (A-Bewertung, »fast«) nahezu die gleiche Lautheit hervor.

Mit Ausnahme der Werte für  $L_{max} = 77\text{dB(A)}$  lieferten die Versuchspersonen reproduzierbare Daten, zum Teil mit geringen individuellen Unterschieden. Im Vergleich zu den vorangegangenen Experimenten erschien jedoch der Lautstärkevergleich von durch Amplitudenbegrenzung verzerrter Sprache und unverzerrter Sprache deutlich schwieriger. Auch geübte Versuchspersonen waren unsicher, ob es ihnen in allen Fällen gelungen war, ausschließlich Lautheitsunterschiede zu skalieren und von den stark unterschiedlichen Klangqualitäten völlig zu abstrahieren.

## 2.5 Fließende Sprache und Sprachrauschen

In dieser Meßreihe verglichen die Versuchspersonen die Lautheit des Test-Satzes mit der Lautheit von sprachsimulierendem Rauschen nach CCITT Rec. G 227. Es handelt sich dabei um ein Breitbandrauschen, das aus Weißem Rauschen

exklusiv Lautheitsdifferenzen und ignorierte vollständig die quite different sound qualities.

## 2.5 Running speech and speech noise

In this test series the examinees compared the loudness of the test sentence with the loudness of speech-simulating noise according to CCITT Rec. G 227. This refers to broadband noise which is derived from white noise by means of a filter with roughly the following attenuation characteristics: no attenuation within the range of 700 Hz, an attenuation of 11 dB at 100 Hz, of 10 dB at 2 kHz and of 40 dB at 8 kHz.

In contrast to the preceding measurements 21 pairs consisting of the test sentence and speech noise of about the same duration were formed. The sound level of the test sentence was kept constant, while the level of speech noise was varied in steps of 2 dB. For the noise we selected level values above and below the maximum speech level. Fig. 5 shows a section of a corresponding level measurement carried out in the library. The A-weighted sound level is represented as a function of time; the pointer deflections of the sound level

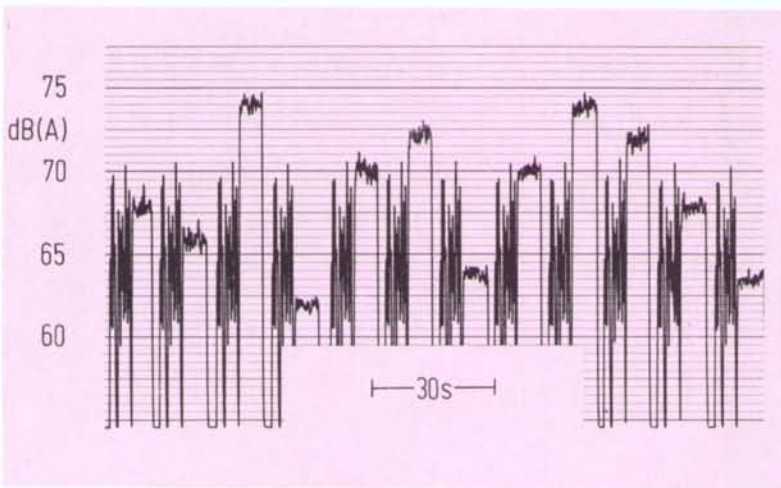


Abb. 5 A-bewertete Schallpegel von fließender Sprache und sprachsimulierendem Rauschen als Funktion der Zeit  
Fig. 5 A-weighted sound levels of running speech and speech-simulating noise as a function of time

mit Hilfe eines Filters etwa folgenden Dämpfungsverlaufes gewonnen wird: keine Dämpfung im Bereich um 700 Hz, bei 100 Hz etwa 11 dB Dämpfung, bei 2 kHz 10 dB und bei 8 kHz 40 dB Dämpfung.

Im Gegensatz zu den vorhergehenden Messungen wurden 21 Paare, bestehend aus Test-Satz und etwa gleich langem Sprachrauschen, gebildet. Der Schallpegel des Test-Satzes wurde festgehalten, der Pegel des Sprachrauschens in 2 dB-Schritten variiert. Für das Rauschen wurden Pegelwerte oberhalb wie auch unterhalb des maximalen Sprachpegels gewählt. Abb. 5 zeigt einen Ausschnitt aus einer entsprechenden, in der Bibliothek durchgeführten Pegelmessung. Der A-bewertete Schallpegel ist als Funktion der Zeit dargestellt; die Zeigeraussschläge des Anzeigeinstrumentes bei schneller Anzeige („fast“) wurden mit einem Pegelschreiber (DC-Eingang) aufgezeichnet.

Aus der in Abb. 5 dargestellten Pegelregistrierung kann entnommen werden, daß der Schallpegel des Test-Satzes zeitliche Schwankungen um mehr als 10 dB aufweist, während für das Sprachrauschen Pegelschwankungen von lediglich etwa 1 dB auftreten. Dem Sprachrauschen kann folglich ein definierter Pegelwert zugeordnet werden, der fließenden Sprache sind jedoch je nach Ablesezeitpunkt deutlich verschiedene Pegelwerte zuzuordnen. Damit stellt sich die Frage, ob für die subjektiv wahrgenommene Lautstärke (Lautheit) fließender Sprache eher das Minimum, ein Mittelwert oder das Maximum des zeitlich schwankenden Schallpegels maßgebend ist.

Eine Antwort auf diese Frage ist aus subjektiven Lautheitsvergleichen von fließender Sprache mit Sprachrauschen zu erwarten. Eine entsprechende Meßreihe wurde folgendermaßen durchgeführt: Wie in den vorangegangenen Experimen-

meter („fast“ display) were registered with a level recorder (DC-input).

From the level recording shown in Fig. 5 we can gather that the sound level of the test sentence is subject to temporal fluctuations of more than 10 dB, while the speech noise level varies only 1 dB approximately. Consequently, a defined level value can be correlated to the speech noise, whereas distinctly different level values are to be correlated to the running speech, depending on the moment of reading. This gives rise to the question whether the subjectively perceived loudness of running speech is decisively influenced by the minimum value, the mean or rather by the maximum value of the temporally fluctuating sound level.

An answer to this question is to be expected from the subjective loudness comparisons of running speech with speech noise. A corresponding test series was carried out in the following manner: As in the previous experiments the value  $N_{ref} = 100$  was again correlated to the loudness of the test sentence; the loudness of the noise had to be assessed accordingly (half loudness  $N = 50$ , equal loudness  $N = 100$  and so on). If from this test series the value  $N/N_{ref} = 100$  per cent is obtained at a speech noise level of 62 dB(A), then it is rather the minimum sound level value which determines the subjectively perceived loudness of the test sentence. By contrast, if  $N/N_{ref} = 100$  per cent at a noise level of 70 dB(A), then it is rather the maximum value of the temporally fluctuating sound level of running speech that is decisive for its loudness (see Fig. 5).

The results of 336 loudness comparison tests which were carried out with 16 examinees, are shown in Fig. 6. The relative loudness  $N/N_{ref}$  of the speech noise is represented as a function of its A-weighted sound level  $L_{SN}$ . The symbols



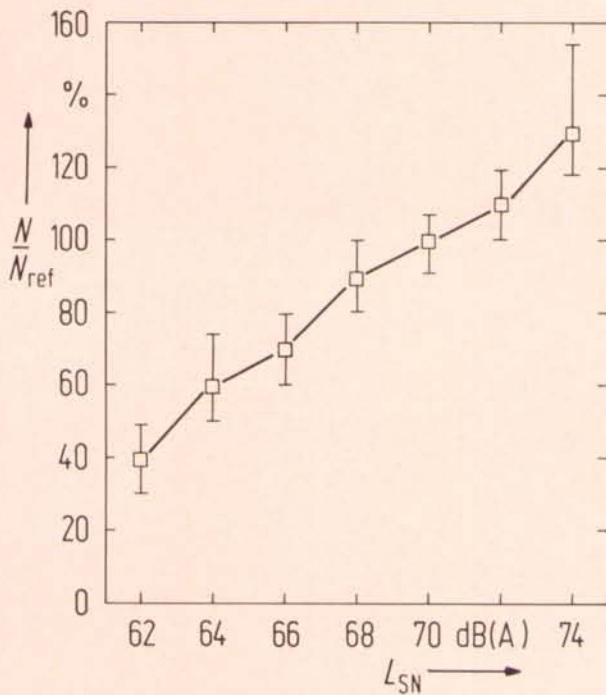


Abb. 6 Lautheitsvergleich von fließender Sprache und sprachsimulierendem Rauschen nach CCITT. Relative Lautheit  $N/N_{ref}$  des Sprachrauschens als Funktion seines A-bewerteten Schallpegels  $L_{SN}$

Fig. 6 Loudness comparison between running speech and speech-simulating noise in accordance with CCITT. Relative loudness  $N/N_{ref}$  of the speech noise as a function of its A-weighted sound level  $L_{SN}$

ten wurde der Lautheit des Test-Satzes wieder der Zahlenwert  $N_{ref} = 100$  zugeordnet; die Lautheit des Rauschens war entsprechend zu skalieren (halbe Lautheit  $N = 50$ , gleiche Lautheit  $N = 100$  usw.). Ergibt sich aus dieser Meßreihe der Wert  $N/N_{ref} = 100$  Prozent bei einem Schallpegel des Sprachrauschens von 62 dB(A), so ist für die subjektiv wahrgenommene Lautstärke (Lautheit) des Test-Satzes eher dessen Schallpegelminimum maßgebend. Zeigt sich dagegen  $N/N_{ref} = 100$  Prozent bei einem Pegel des Rauschens von 70 dB(A), dann ist eher das Maximum des zeitlich schwankenden Schallpegels fließender Sprache für deren Lautheit maßgebend (vgl. Abb. 5).

Ergebnisse von 336 Lautheitsvergleichen, die 16 Versuchspersonen durchführten, zeigt Abb. 6. Die relative Lautheit  $N/N_{ref}$  des Sprachrauschens ist als Funktion seines A-bewerteten Schallpegels  $L_{SN}$  dargestellt. Die Symbole repräsentieren Zentralwerte mit wahrscheinlichen Schwankungen, abgeleitet aus jeweils 48 Meßpunkten.

Für gleiche Lautheit von Test-Satz und Sprachrauschen ( $N/N_{ref} = 100$  Prozent) ergibt sich gemäß Abb. 6 ein Schallpegel des sprachsimulierenden Rauschens von 70 dB(A). Bei den Schallpegeln  $L_{SN} = 68$  dB(A) bzw. 72 dB(A) zeigen sich Werte von  $N/N_{ref} = 90$  Prozent bzw. 110 Prozent; angesichts der in Abb. 6 dargestellten wahrscheinlichen Schwankungen, müssen Sprachrauschen mit diesen Schallpegeln als subjektiv nahezu ebenso laut wie der Test-Satz angesehen werden. Die einzelnen Versuchspersonen reproduzierten ihre Werte gut; nur bei  $L_{SN} = 74$  dB(A) zeigten sich nennenswerte individuelle Unterschiede.

Für die Lautheit fließender Sprache ist demnach offenbar das Maximum des A-bewerteten Sprachschallpegels von ent-

denote medians with interquartile ranges each derived from 48 data points.

If the loudness of the test sentence and the speech noise is equal ( $N/N_{ref} = 100$  per cent), a sound level of 70 dB(A) is obtained for the speech-simulating noise (Fig. 6). For the sound levels  $L_{SN} = 68$  dB(A) or 72 dB(A) values of  $N/N_{ref} = 90$  per cent and 110 per cent are obtained. In view of the interquartile ranges displayed in Fig. 6 speech noise with these levels must be regarded as being subjectively nearly as loud as the test sentence. The individual examinees reproduced their values with good results; considerable differences from person to person were only seen at the sound level  $L_{SN} = 74$  dB(A).

Consequently, the maximum value of the A-weighted speech-sound level appears to be of decisive importance for the loudness of running speech (cf. [10]). With respect to practical applications this result can be interpreted as follows: if, for instance, for objective measurements, running speech is to be replaced by a signal with a temporally constant level, speech-simulating noise appears to serve this purpose well (see [5], [6]). If, in addition, the perception of loudness produced by running speech is to be simulated approximately, then the speech noise level to be selected must correspond to the maximum A-weighted sound level of running speech.

### 3. Conclusions

In all experiments, one and the same test sentence from a reference tape was used. The test material represents a typical, practice-oriented example of running speech (newscaster) that is readily available for further tests at any time and place.

scheidender Bedeutung (vgl. [10]). Dieses Ergebnis läßt sich im Hinblick auf praktische Anwendungen folgendermaßen deuten: Soll zum Beispiel für objektive Messungen fließende Sprache durch ein Signal mit zeitlich konstantem Pegel ersetzt werden, so erscheint sprachsimulierendes Rauschen dafür gut geeignet (vgl. [5], [6]). Soll ferner die Lautstärkeempfindung, die fließende Sprache erzeugt, näherungsweise nachgebildet werden, so ist für das Sprachrauschen ein Schallpegel zu wählen, der dem maximalen A-bewerteten Schallpegel fließender Sprache entspricht.

### 3. Schlußbemerkungen

In sämtlichen Versuchen wurde ein und derselbe Test-Satz aus einem Bezugsband verwendet. Er stellt ein typisches, an der Praxis orientiertes Beispiel für fließende Sprache dar (»Nachrichtensprecher«) und ist für weitere Untersuchungen jederzeit überall greifbar. Darüber hinaus erleichtert die wiederholte Verwendung desselben Textes den Versuchspersonen die während der Lautheitsvergleiche unumgänglich notwendige Abstrahierung von sprachlicher Information, Sprachverständlichkeit usw.

Im folgenden sollen die wesentlichen, sich aus obigen Meßreihen ergebenden Anhaltswerte zur Lautheit fließender Sprache zusammengefaßt werden:

a) Wird fließende Sprache in einem normalen Raum mit Zimmerlautstärke dargeboten, so halbiert sich die Lautheit bei einer Pegelerniedrigung um 7 dB.

b) Werden bei der Übertragung fließender Sprache die hohen Frequenzanteile stark gedämpft, so ergibt sich dadurch kaum eine Lautheitsabnahme. Ausgehend von einer Breitbandübertragung (16 kHz) entspricht die Verringerung der Bandbreite um etwa 95 Prozent ( $f_g = 700$  Hz) subjektiv nur einer Halbierung der Lautheit fließender Sprache.

c) Eine Dämpfung der tiefen Frequenzanteile fließender Sprache führt zu deutlichen Lautheitseinbußen. Bezogen auf eine Breitbandübertragung (16 kHz) ruft eine nur unwesentliche Bandbreitenverringernng um etwa 5 Prozent ( $f_g = 1000$  Hz) bereits eine Halbierung der Lautheit fließender Sprache hervor.

d) Wird fließende Sprache durch Amplitudenbegrenzung verzerrt, zeigen sich auch bei signifikanter Übersteuerung (15 dB) nur geringe Lautheitsunterschiede. Verzerrte und unverzerrte Sprache werden bei gleichem maximalem Schallpegel als nahezu gleich laut wahrgenommen.

e) Die fließender Sprache zugeordnete Lautheit kann meßtechnisch näherungsweise folgendermaßen erfaßt werden: die objektiven Messungen sind mit sprachsimulierendem Rauschen durchzuführen, das einen Schallpegel aufweist, der dem maximalen Schallpegel der Sprache entspricht.

Furthermore, the repeated use of the same text makes it easier for the examinee to ignore the linguistic information, speech intelligibility and so on; this is imperative for the loudness comparison test.

In the following, the essential guide values for the loudness of running speech which were derived from the above-mentioned test series are summarized:

a) If running speech is presented in a normal room at domestic loudness level, the loudness is halved with a level decrease of 7 dB.

b) If during a transmission of running speech, the high-frequency components are strongly attenuated, the loudness is nevertheless hardly diminished. Based on broadband transmission (16 kHz) the limitation of the bandwidth by about 95 per cent ( $f_g = 700$  Hz) corresponds subjectively only to a 50 per cent decrease in the loudness of running speech.

c) An attenuation of the low-frequency components of running speech leads to a significant loss of loudness. Referred to a broadband transmission (16 kHz) a negligible limitation of the bandwidth by about 5 per cent ( $f_g = 1000$  Hz) already halves the loudness of running speech.

d) If running speech is distorted by peak clipping even significant clipping (15 dB) produces only small loudness differences. Distorted and undistorted speech are perceived as being nearly equally loud for equal maximal sound level.

e) The loudness correlated to running speech can be measured approximately in the following manner: The objective measurements are to be carried out with speech-simulating noise having a level that corresponds to the maximum sound level of speech.

### Literatur/References

- [1] Zwicker, E.: Der Einfluß der zeitlichen Struktur von Tönen auf die Addition von Teillautheiten. *Acustica* 21 (1969), 16.
- [2] Fastl, H.: Loudness and masking patterns of narrow noise bands. *Acustica* 33 (1975), 266.
- [3] Proc. FASE '75 (1975), Thema II: Lautheit der Sprach- und Musiksignale.
- [4] Mendel, M. I. et al.: Loudness judgements of speech and nonspeech stimuli. *J. Acoust. Soc. Amer.* 46 (1969), 1556.
- [5] Brinkmann, K.: Die Neuaufnahme des Marburger Satzverständnistestes. *Z. Hörger. Akustik* 13 (1974), 190.
- [6] Fastl, H.: Schallpegel und Lautstärke von Sprache. *Acustica* 35 (1976), 341.
- [7] Zwicker, E. und Feldtkeller, R.: Das Ohr als Nachrichtenempfänger. 2. erw. Aufl., Hirzel Verlag Stuttgart (1967).
- [8] Pollack, I.: On the measurement of the loudness of speech. *J. Acoust. Soc. Amer.* 24 (1952), 323.
- [9] Gleiss, N.: The loudness of connected speech. in: *Speech communication* (G. Fant, Ed.) Almqvist & Wiksell, Stockholm 1975.
- [10] Pedersen, O. J. and Pederson, S. B.: Possibilities for the objective measurement of the loudness level of speech. *Proc. 6. ICA, Tokyo* 1968, A-65.