

TECHNICAL REVIEW

No. 3 — 1976

Contents

Registration of Voice Quality by B. Frøkjær-Jensen and S. Prytz	3
Acoustic Response Measurements and Standards for Motion-Picture Theatres by E. Rasmussen	18

Registration of Voice Quality¹

by

*Børge Frøkjær-Jensen² and
Svend Prytz³*

ABSTRACT

Long-time-average-spectra recordings of normal voices as well as an average spectrum of such LTAS-registrations are shown and discussed.

For comparisons of voice qualities we have tried to set up a new parameter, α , which is a measure of the intensity relations in the higher and the lower parts of the speech spectrum:

$$\alpha = \frac{\text{intensity above 1000 Hz}}{\text{intensity below 1000 Hz}}$$

Because the spectrum above 1000 Hz is normalized relative to the spectrum below 1000 Hz, α is independent of microphone distance, amplification level, etc. α seems to be a good acoustic correlate to the physiological term "medial compression", and preliminary research indicates that it is relevant to evaluations and comparisons of voice qualities.

The "quality parameter" is represented graphically by histograms showing the number of α -values automatically sampled during a read text, and it is displayed on a storage oscilloscope along the vertical axis. The horizontal axis is used for displaying the total speech intensity.

SOMMAIRE

Des spectres enregistrés avec de longs temps d'intégration pour des voix normales et un spectre représentant la moyenne de tels spectres sont présentés et discutés.

Pour comparer les qualités des voix, on a proposé l'emploi d'un nouveau paramètre, α , qui donne une mesure des relations d'intensité entre les parties inférieures et supérieures du spectre de la parole:

$$\alpha = \frac{\text{intensité au-dessus de 1000 Hz}}{\text{intensité au-dessous de 1000 Hz}}$$

1) Paper presented at the International Congress of Phonetic Sciences in Leeds, August 1975. In the present version, slight modifications have been made.

2) Audiologopedic Research Group, University of Copenhagen, 96 Njalsgade, 2300 Copenhagen S, Denmark.

3) Svend Prytz, M.D., Phoniatic Laboratory, ENT-department, University Hospital, Blegdamsvej, 2100 Copenhagen Ø, Denmark.

Comme le spectre au-dessus de 1000 Hz est normalisé par rapport au spectre au-dessous de 1000 Hz, α ne dépend pas de la distance au microphone, du gain, etc. α semble être un bon équivalent acoustique à la "compression médiale" physiologique, et les recherches préliminaires indiquent que α est bien adapté aux évaluations et comparaisons des qualités des voix.

Le "paramètre qualité" est représenté graphiquement par des histogrammes montrant le nombre de valeurs de α automatiquement échantillonnées pendant la lecture d'un texte, et il est affiché sur un oscilloscope à mémoire selon l'axe vertical. L'axe horizontal sert à représenter l'intensité totale du discours.

ZUSAMMENFASSUNG

Es werden die Aufzeichnungen von über lange Zeit gemittelter Spektren normaler Stimmen, als auch ein gemittelttes Spektrum über diese Aufzeichnungen gezeigt und diskutiert.

Zum Qualitätsvergleich von Stimmen haben wir versucht, einen neuen Parameter α einzuführen. Dieser stellt eine Maßzahl für das Intensitätsverhältnis zwischen den höherfrequenten und niederfrequenten Anteilen des Sprachspektrums dar.

$$\alpha = \frac{\text{Intensität oberhalb 1000 Hz}}{\text{Intensität unterhalb 1000 Hz}}$$

Da das Spektrum oberhalb 1000 Hz relativ zum Spektrum unterhalb 1000 Hz normalisiert wird, ist α vom Mikrofonabstand, Verstärkungspegel u.s.w. unabhängig. α scheint eine akustische Größe zu sein, die gut mit dem physiologischen Ausdruck "mittlere Kompression" übereinstimmt. Voruntersuchungen zeigten, daß dieser Parameter für Ermittlungen und Vergleiche von Stimmqualitäten geeignet ist.

Der "Qualitätsparameter" wird graphisch durch Histogramme repräsentiert, welche die Anzahl von α -Werten, die automatisch während des Lesens eines Textes gewonnen werden, aufzeigen. Er wird längs der vertikalen Achse auf einem Speicheroszillogoskop angezeigt. Die horizontale Achse wird zur Anzeige der gesamten Sprachintensität benützt.

Introduction

We define voice quality as an auditory property, i.e. an aspect of the perception of the human voice. A good voice quality depends on (1) certain typical formant patterns, (2) absence of noise in the acoustic spectrum, and (3) a high degree of absence of aperiodicity of the fundamental frequency.

In 1963 Wendahl and Moore found a direct relation between the jarring, rough, and hoarse voice quality in voices suffering from unilateral internus paralysis and the variations in periodicity between adjacent pitch periods.

Lieberman has defined these pitch variations in terms of the so-called Lieberman pitch perturbation factor, and he has analyzed the magnitude of this factor in different larynx disorders by computer.

Smith and Lieberman found significant differences between normal subjects and patients suffering from cancer of the vocal folds, polyps on or adjacent to the vocal folds, and acute and chronic laryngitis.

Koike improved this method and obtained similar results, whereas Hecker and Kreul could not reproduce Lieberman's results, even though the methods were almost identical. They defined instead a "directional pitch perturbation factor", which depends on the direction of the perturbation change. This factor was a significant improvement in the discrimination of pathological from healthy voices. Furthermore, they found a more narrow distribution of the fundamental frequency in pathological voices than in normal voices, and they established that the averaged fundamental frequency and duration of phonations were reduced compared to the normal voices. However, the investigation patients were all selected and matched, and it was found that they all suffered from laryngeal cancer.

Hans von Leden and Iwata have investigated pitch perturbations (among other diseases) in 10 patients of unilateral recurrent paralysis before and after teflon[®]-injection in the paralysed vocal fold. They found the method reliable and useful in the phoniatic clinic. The discrimination between different laryngeal diseases was poor, however.

Just as important as the cycle-to-cycle variations in pitch is the acoustic structure of the speech spectrum. According to the literature, this spectrum has its origin in the voice source and decreases by 12 dB per octave. However, during normal speech we find variations in the slope. Glottography and inverse filterings seem to show that the slope for voiced consonants is about -15 dB per octave, and thus steeper than the slope for vowels. On the other hand, we find changes in the opposite direction during high voice effort, such as shouting.

Changes in voice quality are used by singers and actors as an artistic way of expressing their emotions and moods, whereas vocal disability as well as voice disorders create unpleasant voice qualities, such as breathiness, hoarseness, and roughness.

Within the phoniatic and logopedic clinic there is a great demand for developing instrumentation and methods for registration of changes in voice quality.

The present paper is a preliminary report, dealing with three different methods for voice quality analyses:

- (1) Long-time-average-spectral analysis based on a read text of a duration of 45 seconds.
- (2) Histograms of the voiced part of speech showing the amplitude level above 1000 Hz, relative to the level below 1000 Hz.
- (3) The relative amplitude parameter shown as a function of the total amplitude level on a storage oscilloscope screen.

To our knowledge, long-time-average-spectra of human speech have only been employed sporadically in the last 25 years, partly due to the sophisticated instrumentation needed for these analyses, especially when performed in real time. Only a few laboratories have been able to set up their own LTAS-analyzing system.

Measurement Procedure

The long-time-average-spectra were obtained using a 400 channel Brüel & Kjær Real Time Narrow Band Analyzer Type 3348 which consists of a spectrum analyzer, an averager and a 12" CRT display unit. The built-in Hanning weighting network was used to obtain a better selectivity giving a 3 dB bandwidth of 18,75 Hz and a resolution of 12,5 Hz. A full spectrum analysis was carried out every 80 ms in the frequency range 0 — 5000 Hz.

In the averager the spectra have been averaged over a period of 45 seconds with a 50 dB dynamic range of the resulting average spectrum. On account of pauses in the read text, which effectively lowers the average level, a 6 dB/octave shaping filter in the frequency range 100 — 5000 Hz was introduced, Fig.1. Furthermore, a gating system that cuts out all voiceless speech segments was also included. A "post averaging gain" of + 12 dB, incorporated in the averager, was used before displaying the average spectrum.

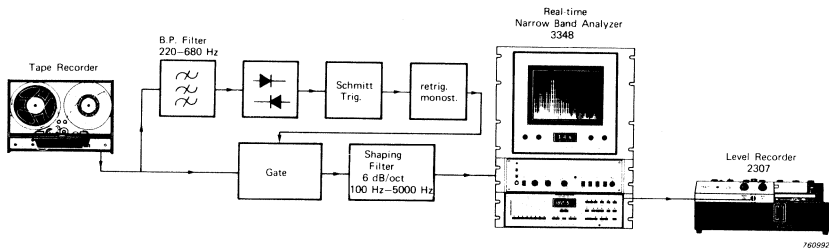


Fig. 1. Measuring Arrangement

Two digital memories are built-in the averager for storing spectra. The contents of either memory may then be continuously displayed or the contents of both may be displayed alternately. Hence comparisons of LTAS before and after speech therapy can easily be made.

All of the 400 channels are updated every 45 ms in the display unit which also contains analogue and digital output facilities. For obtaining hard copy of the results a Level Recorder Type 2307 is included in the set-up.

Investigations were carried out in order to test the validity of LTAS analyses as an objective means for the evaluation of the spectral changes that occur during the recovery period in patients suffering from speech disorders.

Typically, the LTAS analyses of recurrent paralysis may be divided into two groups, based upon the spectral changes:

- (1) Patients who get a *reduced* energy spectrum above the first formant region during the recovery, and
- (2) patients who get an *increased* energy spectrum above the first formant region during the recovery.

Both groups consist of patients having a dyscoordination between the subglottal air pressure and the medial compression of the vocal folds. In the first group the result is a phonatory hyperfunction which emphasizes the higher part of the spectrum and diminishes the stability of the fundamental frequency (rough voice quality), often to such a degree that diplophonia occurs.

The second group, on the contrary, is characterized by a phonatory hypofunction where the vocal folds cannot close very well, which results in an acoustic spectrum with weak higher harmonics. Often some breathy noise is heard, caused by turbulent air passing through the glottis.

Fig.2 and Fig.3 are illustrations of the two types of changes in the LTAS. In Fig.2 (female voice suffering from recurrent paralysis recorded before and during speech therapy) we observe that the acoustic spectrum below 1000Hz has not been changed very much during therapy. Above 1000Hz, however, the higher harmonics of the spectrum are weaker than before treatment. The three peaks in the first formant region depict the first three harmonics. If the voice is a good one with

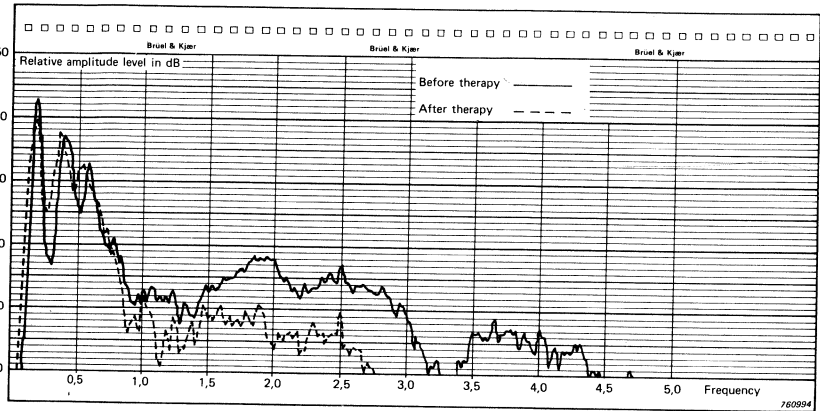


Fig. 2. Spectral distribution of speech amplitude level averaged over 45 seconds. Female patient with unilateral recurrent paralysis

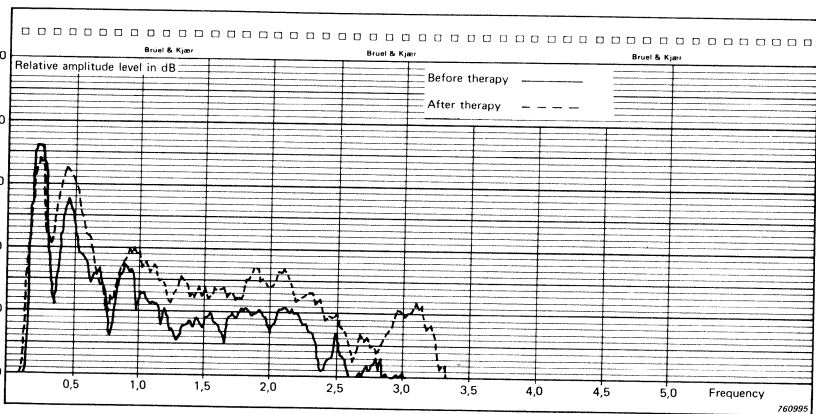
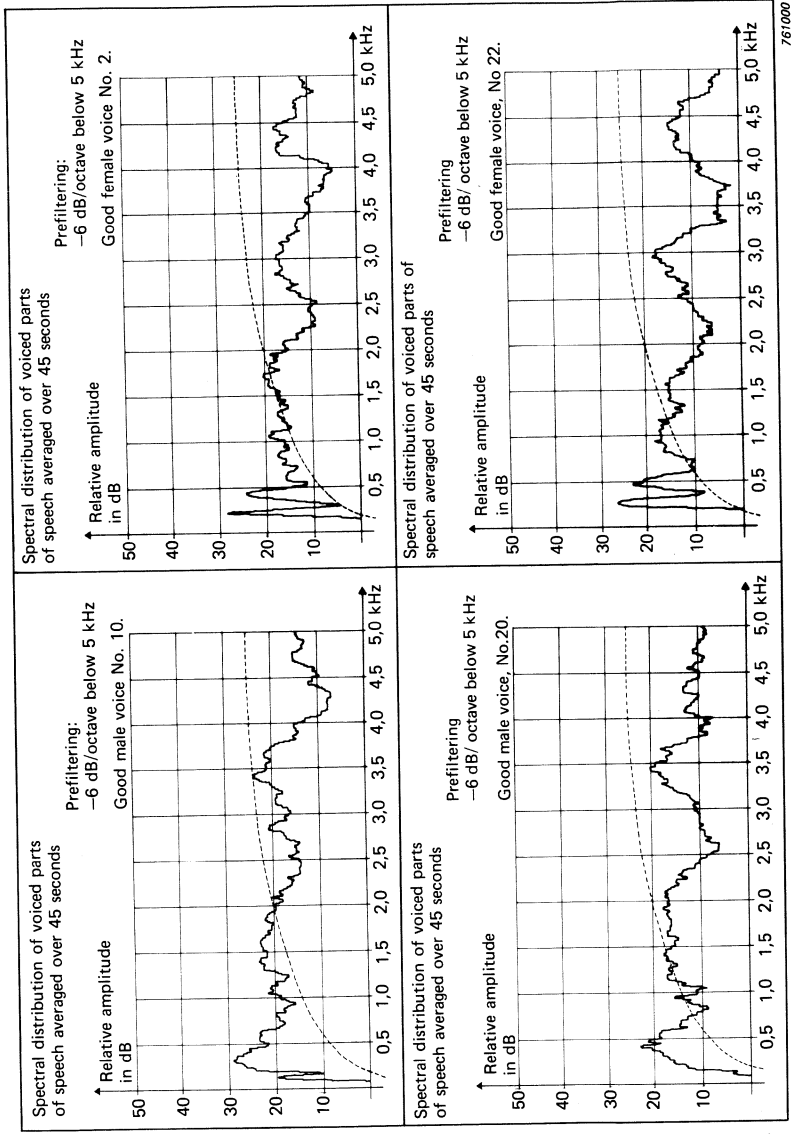


Fig. 3. Spectral distribution of speech amplitude level averaged over 45 seconds. Female patient with unilateral recurrent paralysis

pitch variations over a wide frequency range, this discontinuity ought not to appear, but for unhealthy voices we observe a pronounced harmonic pattern caused by lack of pitch modulations. Above the first formant region we normally notice some spectral discontinuity caused by the average levels of the second and third formant. Between 3,5 and 4,5 kHz we find a broad energy maximum without any importance for the intelligibility. During speech therapy this energy maximum has



761/000

Fig. 4. LTAS-analyses of good and healthy voices

been diminished with the auditory result that the voice sounds less shrilly.

Fig. 3 shows a case where speech therapy causes an increase in the spectral energy (in this case about 4 dB) except for the first harmonics, the levels of which are reduced a few dB. This is a common case of normal therapy progress for recurrent paralysis.

Fig. 4 shows four LTAS-analyses taken from 22 normal voices. Along the Y-axis we have the relative amplitude level in dB. The dotted line indicates the above-mentioned preemphasis of +6 dB per octave. We observe deviations among the four voices. Especially for voices Nos. 2 and 22 we observe a pronounced depiction of the lower harmonics, which we may interpret as restricted variations in the fundamental frequency or intonation for these two voices. It may be due to the subjects' behaviour during the recording procedure. We do not find this harmonic pattern in voices Nos. 10 and 20.

Fig. 5 shows the spectral distribution of 10 normal voices set up in the same graph. Notice the relatively small spread among the curves, which indicates that the spectral distribution of normal and healthy voices is fairly constant, at least up to about 3000 Hz.

Spectral Distribution of Voiced Parts
of Speech Averaged over 45 seconds

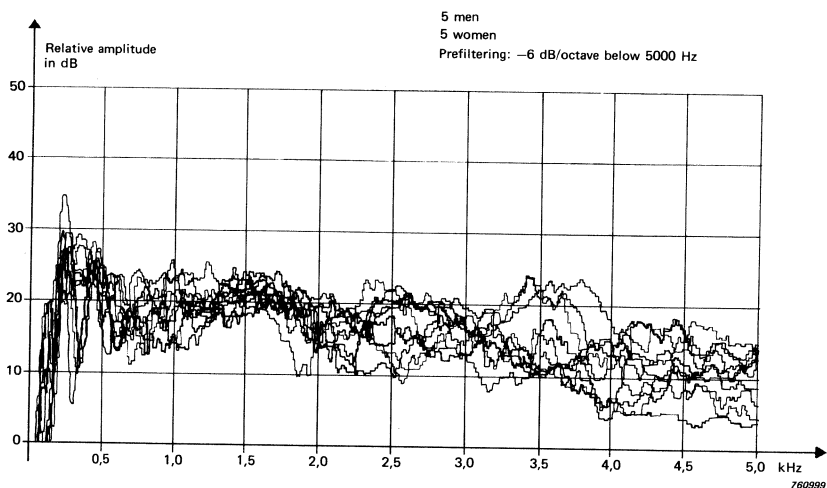


Fig. 5. Long-time-average-spectra of 10 normal voices

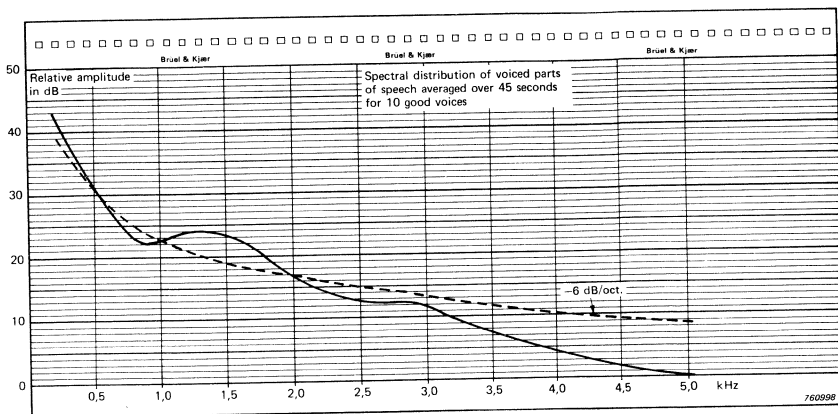


Fig.6. The mean curve of the 10 spectra shown in Fig.5 without pre-filtering

In Fig.6 the solid line shows the average curve of the 10 curves shown in Fig.5, but without the prefiltering of 6 dB/oct below 5000 Hz. The dotted line indicates the commonly presumed slope of -6 dB per octave of the radiated sound wave (voice source + radiation). As may be expected, we notice that the slope of the averaged speech spectra is steeper than -6 dB per octave.

Just around 800 Hz we find a minimum in the radiated sound spectrum, but we cannot, based upon these recordings, decide whether this is due to less frequent occurrence of formant energy around this frequency, or whether it is due to a zero in the voice source.

In the following a method is shown for quantifying the difference in the spectra of voices suffering from unilateral recurrent paralysis before and after therapy — not for the purpose of showing what happens during the treatment of a given disorder, but merely to show how these analyses could be used for comparisons of the voice qualities.

For these comparisons we have tried to set up a new parameter, which we have called α .

We have defined

$$\alpha = \frac{\text{amplitude level above 1000 Hz}}{\text{amplitude level below 1000 Hz}}$$

$$\log \alpha = \log A \text{ (above 1000 Hz)} - \log A \text{ (below 1000 Hz)}$$

Because the amplitude above 1000 Hz is normalized relative to the amplitude below 1000 Hz, α is independent of the microphone distance, amplitude levels, etc.

Fig.7 shows how the α -parameter is extracted from the tape recordings. In a differential amplifier we get the difference between the logarithmic voltages proportional to the intensity levels above and below 1000 Hz. It does not matter whether we use intensities or amplitude levels, it will only be a question of calibration, because the intensities are proportional to the square of the sound pressure level.

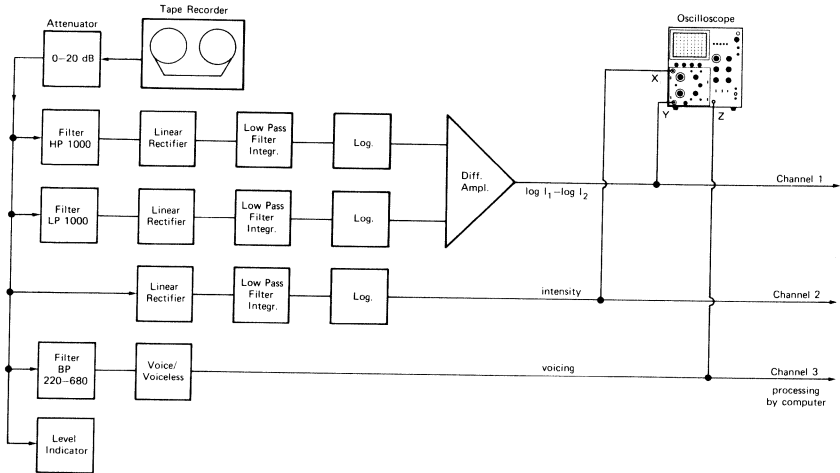


Fig. 7. Instrumental set-up for registration of: total amplitude level, difference between amplitude levels above and below 1000 Hz, and duration

The set-up includes a voice/voiceless indicator based upon a sensing of the energy in the first formant region, and a full frequency logarithmic intensity channel.

The α -parameter is displayed on a storage oscilloscope as a function of the total intensity, where the light intensity of the oscilloscope is switched off and on by the voice/voiceless indicator.

α may also be recorded automatically 25 times per second and represented as a histogram by the computer.

Fig.8 shows the LTAS-analysis of Fig.3 and α -histograms of a patient with a phonatory hypofunction caused by recurrent paralysis before and

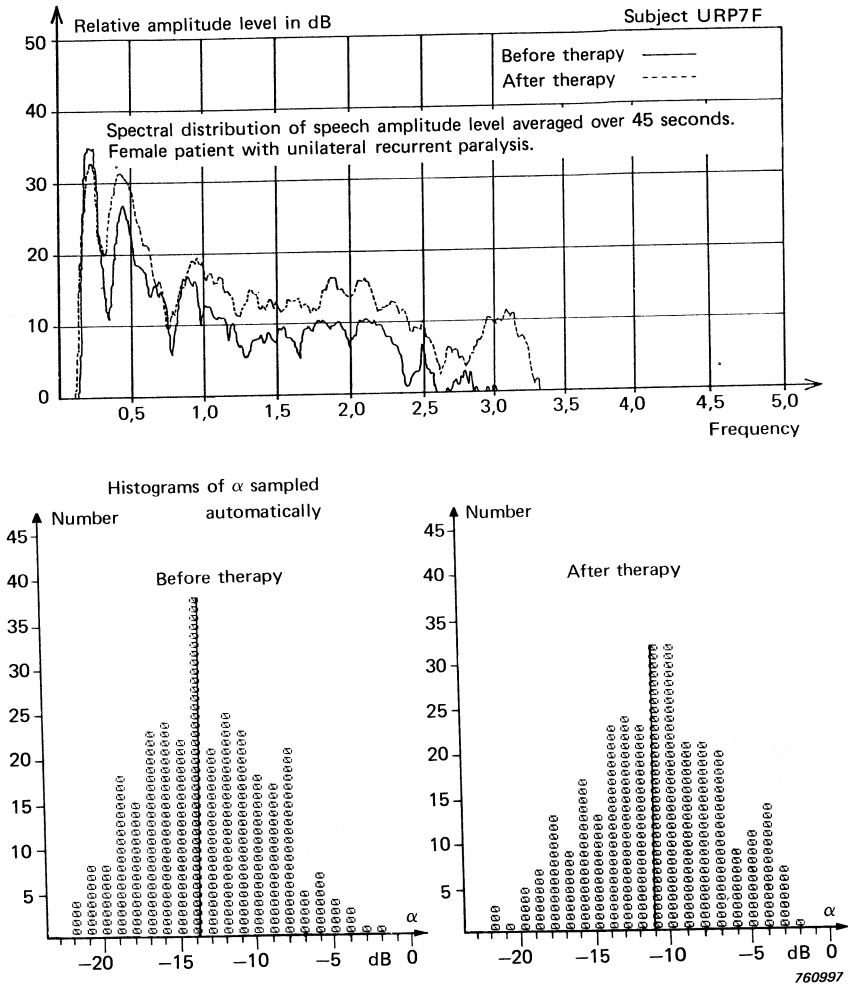


Fig.8. Upper graph: LTAS-analysis of a patient suffering from unilateral paralysis before and after treatment
Lower graphs: Distribution of the amplitude above 1000 Hz relative to the amplitude below 1000 Hz, sampled automatically 25 times per second

after speech therapy. The graph shows how much the spectral amplitude has increased at different frequencies in the spectrum during treatment. The LTAS-graph is typical in the sense that the energy is increased except for the fundamental frequency which has been weakened.

Examinations of LTAS-graphs from the voices of more than 50 patients and several normal subjects reveal that 1000 Hz seems to be a reasonable cut-off frequency for the above-mentioned comparisons between the higher and the lower part of the spectrum. This is in agreement with Ilse Lehiste, Gordon Peterson and Svend Smith.

The histograms of α before and after treatment in this illustration show an increase of about 4 dB for α .

In the next illustration, Fig.9, we notice an increase of about 3 dB during the speech training.

Fig.7 above showed the instrumental set-up for recording the α -parameter. As it appears from that illustration, the α -parameter could also be shown on an oscilloscope as a function of the total intensity. This is il-

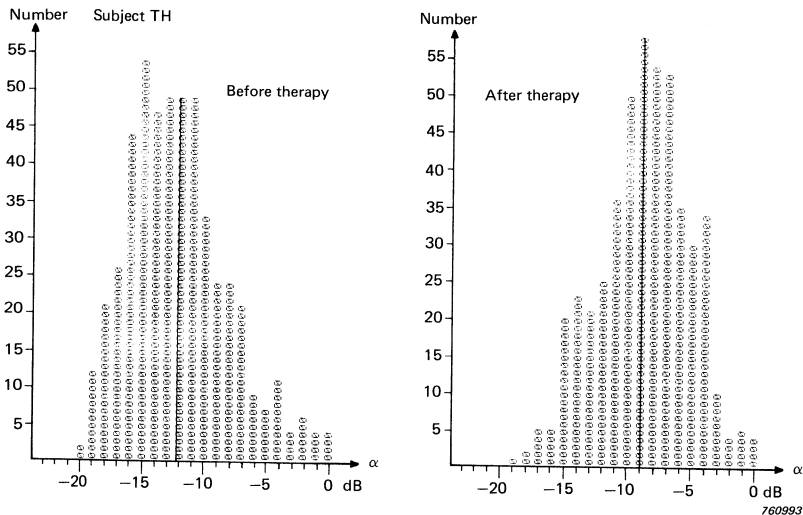
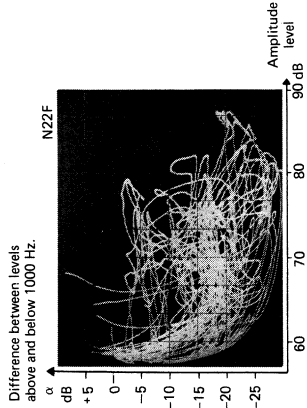
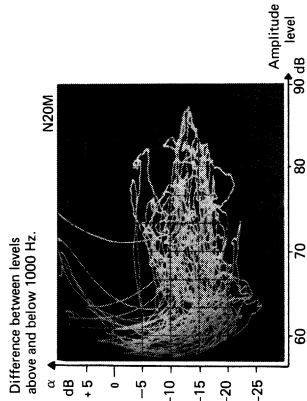
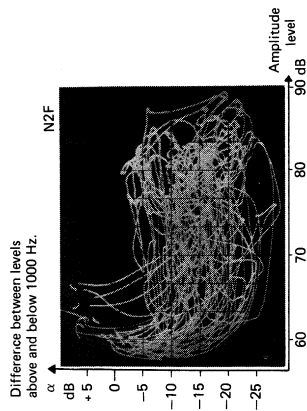


Fig.9. Sample of α -graphs before and after speech training, showing an increase of about 3 dB of the spectral amplitude above 1000 Hz relative to the spectral amplitude below 1000 Hz

Good voice qualities



Pathological voices

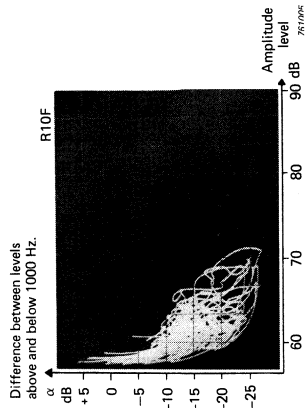
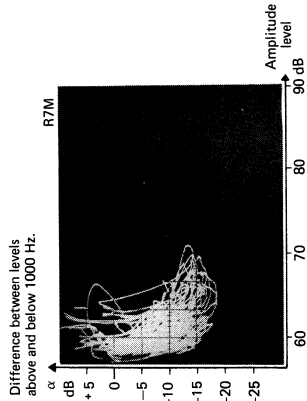
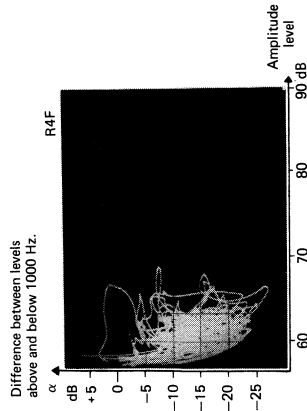


Fig. 10 Oscilloscope displays of total amplitude level (in dB SPL) and relative amplitude above 1000 Hz

illustrated in Fig.10. The photos of the storage screen of the oscilloscope depict the α -parameter as a function of the total intensity, averaged over 20 seconds. In this illustration we have given three healthy and three pathological voices for demonstration purposes only.

The X-axis in these photos is calibrated in dB sound pressure level, measured at a distance of 30 cm from the mouth. The Y-axis is calibrated by means of known synthetic vowel spectra and attenuated tones. Theoretically we may expect a voice phonating with a high voice effort to be placed in the right-hand part of the photos, and a voice with low voice effort to be depicted in the left-hand part of the photos. Voices with a low α -value will be depicted in the lower part, and voices with a high α -value will be placed in the upper part of the photos.

These differences between normal and pathological voice qualities may be noticed in Fig.10. They are most obvious as regards the total intensity, but also the α -parameter shows mutual differences, e.g. between voices No. R 7M and No. R 10F. In fact, voice No. R 7M sounds as a hyperfunctional dysphonia*, and voice No. R 10F as a weak, breathy voice.

As oscilloscope registrations of this kind are fairly simple to make, it seems that they might be useful in the phoniatic clinic as a quick check of some important characteristics of the voice. Further research may prove this.

Conclusion

We have made a pilot test of some new methods for long-time-averaging of the balance between the lower and higher parts of the speech spectrum. This balance depends on the voice source and seems to correlate with the term "voice quality" which, unfortunately, is still a badly defined term. The registration methods seem to be valid if the acoustic spectrum is not dominated by white noise. Further research with pathological voice qualities produced synthetically may show to what extent the method is valid when applied to very noisy voices.

The coming years may show if registrations of the intensity above 1000 Hz relative to the intensity below 1000 Hz will turn out to be a useful aid in the phoniatic and logopedic routine diagnosis, as well as a tool for voice evaluation during speech therapy.

* Hyperfunctional dysphonia is the medical expression for an unpleasant voice quality caused by an improper balance between subglottal air pressure and too strong muscle tension in the larynx and the neck.

References

1. BUCH, N. H. and
B. FRØKJÆR-JENSEN: "Some Remarks on Acoustic Parameters in Speech Disorders", ARIPUC 6, p.245-259.1972.
2. FRØKJÆR-JENSEN, B.
and S. PRYTZ: "Evaluation of Speech Disorders by Means of 'Long-time-average-spectra'", ARIPUC 8, p.227-237.1974.
3. HECKER, M. H. L. and
E. J. KREUL: "Descriptions of the Speech of Patients with Cancer of the Vocal Folds", JASA 49, 4, p.1275-1282.1971.
4. IWATA, S. and
HANS von LEDEN: "Pitch Perturbations in Normal and Pathological Voices", Folia Phoniatria 22, p.413-424.1970.
5. LEDEN, HANS von and
Y. KOIKE: "Detection of Laryngeal Disease by Computer Technique", Arch.Otolaryng. 91, p.3-10.1970.
6. LEHISTE, I. and
G. PETERSON: "The Identification of Filtered Vowels", Phonetica 4, p.161-177.1959.
7. SMITH, S.: "On Artificial Voice Production", Proc. Phon. IV, Helsinki, p.96-110. 1961.
8. SMITH, W. R. and
P. LIEBERMAN: "Computer Diagnosis of Laryngeal Lesion", Computers & Biomedical Research II, p.291-303.1969.
9. WENDAHL, R. W.: "Some Parameters of Auditory Roughness", Folia Phoniatria 28, p.26-32.1966.
10. WINCKEL, F.: "Darstellung des Sprechverhaltens als Statistische Tonhöhen und Formantverteilung mittels 'Langzeitanalyse'", Proc.Phon. VI, Praha, p.1031-1035.1967.

Acoustic Response Measurements and Standards for Motion-Picture Theatres

by

Erik Rasmussen

ABSTRACT

A proposal of the Nordisk Film/TV Union (NFTU), document 1069—1969, is discussed in great detail. This document refers to the measurement of acoustic response in cinema theatres for optical 35-mm soundtracks. The instructions, procedures and equipment this proposal offers for consideration are described at length, for the benefit of the service engineer for whom this type of measurement may constitute a new field of activity. Also the work carried out simultaneously in other countries and by ISO is described. Reference is made to test films and further developments of instrumentation in the recent past. Mention is made of the new possibilities opened up by an ISO Draft Proposal, calling for an "Y"-curve to cover films currently in production or exploitation, and for a "X"-curve for such new systems as the Dolby SVA and the Hue-Modulated Photographic Color Soundtrack. A hope for further research regarding roll-off or no roll-off is also expressed.

SOMMAIRE

Une proposition de Nordisk Film/TV Union (NFTU), document 1069—1969, est discutée en détails. Ce document traite de la mesure de la réponse acoustique dans les salles de cinéma pour les pistes sonores optiques de 35 mm. Les instructions, procédures et équipements présentés par cette propositions sont largement décrits afin d'aider l'ingénieur de maintenance pour qui ce type de mesure peut constituer en nouveau domaine d'activité. On décrit aussi les travaux effectués simultanément dans d'autres pays et par l'ISO. On fait référence à des films d'essai et aux améliorations apportées à l'instrumentation dans un passé récent. On mentionne les nouvelles possibilités offertes par un projet de norme ISO qui demande l'emploi d'une courbe "Y" pour les films actuellement en production et en exploitation, et d'une courbe "X" pour les nouveaux systèmes tels que le Dolby SVA ou la piste sonore photographique en couleur modulée en teintes. L'espoir de voir mener de nouvelles recherches concernant l'atténuation est aussi exprimé.

This paper was presented on 30th September 1975 at the Society's 117th Technical Conference in Los Angeles (read by Jack Leahy, RCA Corporation) by Erik Rasmussen, 20 Solvænget, DK-2800 Lyngby, Denmark. It is reprinted from the Journal of the Society of Motion Picture and Television Engineers, Vol. 85, No. 3, March 1976, (pp. 164-169).

ZUSAMMENFASSUNG

Ein Vorschlag der norwegischen Film/TV-Union (NFTU), Schriftstück 1069—1969, wird in allen Einzelheiten erörtert. Dieser Artikel befaßt sich mit der Messung von akustischen Übertragungsmaßen in Lichtspieltheatern mit 35 mm Licht-Tonspur. Die Vorschriften, Anweisungen und Meßausrüstungen, die dieser Vorschlag angibt, sind zum Überdenken ausführlich beschrieben. Dies gereicht dem Service-Ingenieur zum Nutzen, für den diese Art von Messungen ein neues Betätigungsfeld eröffnen. Ferner werden auch die Arbeiten, die zu gleicher Zeit bei der ISO ausgeführt werden, beschrieben. Man greift auf Testfilme und Geräteweiterentwicklungen aus jüngster Vergangenheit zurück. Erwähnt werden die neuen Möglichkeiten, die sich durch den ISO-Entwurfsvorschlag eröffnen. Dieser schlägt eine "Y"-Kurve für Filme vor, die noch in der Herstellung oder Benutzung sind sowie eine "X"-Kurve für neuartige Systeme wie Dolby SVA und die farntonmodulierte photographische Farb-Tonspur. Die Hoffnung auf weitere Forschungsarbeiten bezüglich Abfall oder kein Abfall wird auch angesprochen.

Introduction

For years the terms "acoustic response", "electro-acoustic response", and "optical-acoustic response", as well as "listening characteristics", have appeared as titles of papers presented at SMPTE Conferences and published in the *Journal*. Other countries besides the United States have been interested in the same subject, such as the United Kingdom, France, Belgium, Poland and the Scandinavian countries. A small number of technicians are and have been deeply involved in research in this field, but in the author's opinion, there has been some reluctance by the industry to adopt the proposed type of measurement as a routine check for the acoustical characteristic of any cinema, new or old.

Presumably, two reasons for this can be pointed out. First, most of the cinema theatre technicians are trained primarily in electronics. They are familiar with the well defined input-output measurements of all kinds of sound devices as long as sine-wave generators can be used as a signal source. But now we are moving into an entirely new field where we must replace the customary signal source with a pink-noise generator, or pink noise recorded on film or tape. At the same time, and this is the most serious impediment, the output from the acoustic link in the theatre, the loudspeaker, is far from well defined.¹

A second reason for reluctance might be the cost of the indispensable measuring equipment: the sound-level meter and the pink-noise signal source. It is unlikely that the individual theatre owner would be interested in the necessary investment because the tracing of any theatre's "house-curve" is a one-shot situation. Only the replacement of the loudspeaker system or its associated power-amplifier, or a change in the acoustical environment or replacement of the screen, could justify a recalibration.

Another important point is that a single theatre curve is not too interesting. It is, on the other hand, of utmost importance when a greater number of theatres are measured and compared. And last, but not least, such curves are of interest when compared with the studio dubbing-room curve, where the picture is acoustically "tailored" before its final release. The statistical large-scale relation is of utmost importance for the future of this work. It is the foundation stone for all further progress in motion-picture theatrical sound reproduction. The purpose of this paper is to provide a short review of all that has happened in this field since we began our work on the subject in Denmark in 1968, and to reveal some of the impacts and impasses that have delayed the final agreement on an international standard. Also, some practical hints on test equipment will be given.

The NFTU Proposal

The establishment of the Danish Filminstitute (in connection with a new film law) opened up the possibility for a non-commercial investigation of the state of the art of sound reproduction in Danish cinemas. All theatre owners were offered a free-of-charge sound test, and about 50% accepted it. In close collaboration with Sweden, and through a Technical Committee under the Scandinavian Film and TV Society (Nordisk Film/TV Union — NFTU), document 1069—1969 was distributed. Its title is

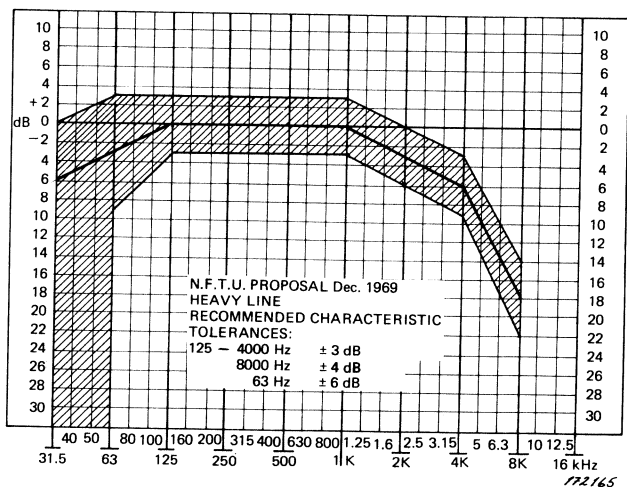


Fig. 1. Optical-acoustic frequency response curve and tolerance limits of NFTU-Proposal 1069—1969

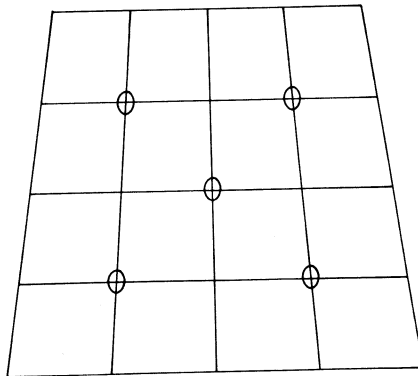
“Recommended Practice for Measuring Overall Frequency Response in Cinemas for Optical Sound 35 mm.”

Its full text, quoted below, gives the following instructions for the service engineer:

“An optical test film, containing 9 bands of octave-filtered pink noise, from 31 to 8000 Hz, is run five times on the (same) projector, and the sound pressure level at each of five measuring points for each of the 9 bands is noted (see Appendix 2 of document 1069). The five readings for each octave band are averaged (see Appendix 3 of document 1069). The resulting 9 values for sound pressure levels constitute a frequency response curve, which shall fit inside the boundaries of Fig.1.”

Preliminary Examination of Theatre Auditorium (Appendix 1 of NFTU 1069)²

“1.1. Acoustical defects. Because these measurements deal only with the steady-state properties of the auditorium, acoustical defects such as back stage overhang, harmful echoes and so on, do not show up. Attempts to use measurement results as a basis for major equipment redesign in a theatre found defective, must be preceded by ascertaining that no grave acoustical faults are present. Methods for finding or eliminating such faults are not covered in this recommendation.



172173

Fig.2. *Determination of measuring points (microphone placements) in accordance with NFTU-Proposal 1069—1969*

1.2. Seating area. A cinema auditorium may have only one seating area — stalls (orchestra) on a single floor. On the other hand, it may have as many as three or more seating areas: stalls in front of the balcony, stalls below the balcony, and one or more balconies. Because the frequency response usually is different in different areas, and because an average taken over all areas may mask useful information, it is recommended that separate measurements be made for each area.

1.3. Selection of measuring points. The seating area is divided into 16 parts, and 5 seats, placed approximately as shown in Figure 2, are chosen as measuring points. Seat position and number are noted. The sound level meter (2.1.2. indicated in the next paragraph) is to be placed at each of these measuring points, in a position representative of a patron's ear height, and with the microphone pointed toward the centre of the screen."

Measuring Equipment³

From Appendix 2, Section 1, we quote:

"2.1.1. Octave pink-noise test film NFTU Type OA 351.

2.1.2. Brüel & Kjær Precision Sound Level Meter, Type 2203 with Microphone 4145, and Octave Filter-set 1613. Note: A floor stand with swivel head is recommended as an accessory.

2.1.3. (Optional) Brüel & Kjær Sound Level Recorder 2307 or

2.1.4. (Optional) DC Millivoltmeter, preferably battery operated, with about 5 sec. integration time."

Measuring Procedure⁴

From Section 2 of Appendix 2, we quote:

"2.2. With the curtain open and all controls for volume and equalization in the position normally used during shows, the test film is run through the (same) projector once for each measuring point, and the sound level in dB for each octave band is noted. During measurement, the octave filter 1613 is switched manually, in accordance with the octave sequence of the test film.

Note 1: Operator position must be such as not to disturb the sound field, in accordance with Brüel & Kjær's recommendations.

Note 2: Levels in the 500-Hz band and upwards can be read directly off the sound level meter dial with an overall accuracy of 1 dB (meter set to SLOW). Lower frequency bands give increasingly lower accuracy at direct read-out, and a sound level recorder, as recommended in 2.1.3., or an integrating meter as per 2.1.4., is of considerable help.

Note 3: In many instances, with low auditorium noise level, it is possible to dispense with the octave filter (2.1.2. mentioned above) and to use the B & K sound level meter directly, set to LINEAR."

Method of Averaging⁵

The following is from Appendix 3:

"The correct method of averaging a number of physical sound pressure levels in a room is to make a conversion from dB, as read out by the instruments indicated above, to sound pressure level p , and take the rms of these pressures and convert back again to dB (ISO Recommendation R. 140)."

Since it is sufficient to know the corresponding figures for dB-level and sound pressure squared, a simple conversion table has been prepared and given in appendix 4⁶ as well as the method of averaging. The same method was also outlined in another article by the author published in Brüel & Kjær Technical Review No.2, 1972.

"*Note:* Averaging dB levels directly, which may be permissible as a quick check, gives good answers only when differences between dB levels are quite small. For differences of 10 dB or more, errors are considerable."

The document was edited in 1969 and at that time we used the 4145 1-in microphone with its random incidence corrector (to avoid increasing response at higher frequencies); today a 4165 1/2-in microphone with the same sensitivity as the 4145 can be obtained and its smaller dimension renders the random incidence corrector superfluous. Any sound-level meter fulfilling the IEC 179 specifications for Precision Sound Level Meters can be used.

It should be mentioned here that the terms sound-pressure level and sound level should not be used at random. *Sound-pressure* level is the actual manometric pressure, expressed commonly in microbars or other manometric units. Sound level, on the other hand, has to do with how this sound is perceived by the ear, i.e., after the insertion of a dB(A), dB(B), dB(C), or dB(D) weighting network.

The NFTU Pink-Noise Test Film

The very first thing we had to do to put proposal 1069 into practice was to organize the production of the pink-noise test film. An idealized diagram of it is shown in Figure 3. As can be seen from this figure, the 8 kHz band has a bracket to indicate that it is a "hybrid" since the theoretical upper limit of 11,3 kHz cannot be reached in optical sound. The band is only about half filled in and this gives a loss of about -3 dB in playback level. To compensate for this, the 8-kHz band is recorded with a $+3$ dB rise to ensure a flat response when played back on a flat system. The SMPTE Multifrequency Test Film Type APFA (1951) has been our reference. (At present, SMPTE 35 mm photographic multifrequency test films carry the code designation P35-MF.) Every specimen of the NFTU pink-noise film OA 351 is a direct positive recording and is accompanied by an individual calibration chart. This film is presently in production.⁷

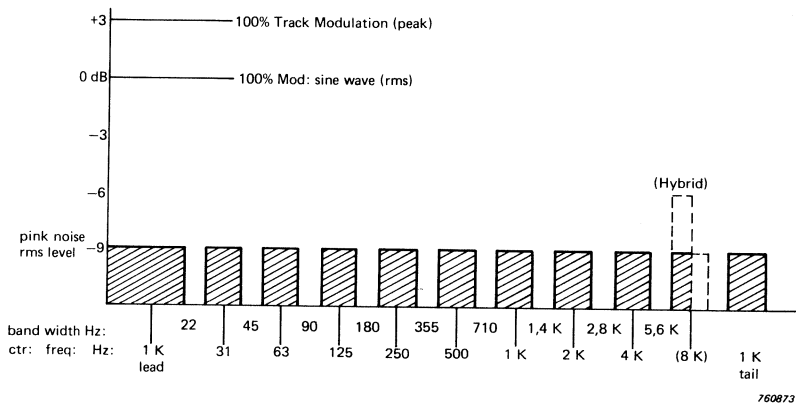


Fig.3. Schematic representation of NFTU octave-filtered, pink-noise test film, type OA 351

The Measurement of Acoustic Response

For greater economy, we considered at first suggesting only the test film and the sound level meter as complete equipment, omitting the octave-filter set. We gave up on this, however, because of a certain danger. When working in a cinema and using the linear scale, traffic noise can give readings as high as 60—65 dB and hardly be noticed by the ear, due to the ear's very limited sensitivity to very low frequencies. For this reason, precisely, our test film has 5 s silent intervals between

the noise bands to enable one to distinguish wanted from unwanted noise. If the "dip" between the bands is filled up with noise, there is something wrong. This could easily be the case when approaching the 8-kHz band which in many cinemas might measure as much as 20—25 dB down compared with the 1-kHz band. Actually, one could compensate for this by switching to the dB(A) scale. A small correction factor could then be added because the dB(A) curve is not completely straight above 1 kHz. We found this too risky, however, and recommend the use of the octave filter.

We have been asked for even further simplification of the acoustic response measurement, namely a single-point measurement in the centre of the theatre. This would certainly reveal some information, but the lower frequency bands could be very difficult to measure in some auditoriums. Also, dispersion of the higher frequencies throughout the auditorium is very important. Hence, we insist on a minimum of 5 measuring points.

There were several good reasons for choosing full octave bands instead of one-third-octave bands. Octave bands agree very well with subjective listening tests. For more thorough investigations, where resonance points and cross-over networks are suspected, one-third-octave bands give a better definition. Narrower bands, however, give a more "ragged" curve. Also, instead of only 9 bands we would need 27 one-third-octave bands on the pink-noise test film. In addition to the 5 measuring points now required, several more would be needed. For a theatre with only one seating area, the number of dB values to be averaged would increase from 45 to several hundred.

Ljungberg⁸ explained that:

"So many factors of importance for the matching of listening with measuring results have been left out in these investigations, that third octave instead of full octaves are felt at this point as a refinement in the wrong direction. If octave results do not agree with listening impressions, much more important avenues are reverberation time, reflection patterns and transient behaviour.

A third-octave measurement in more or less anechoic conditions, to check speaker details, plus octave measurements in the auditorium, to check listening balance, is a good combination, though not always feasible."

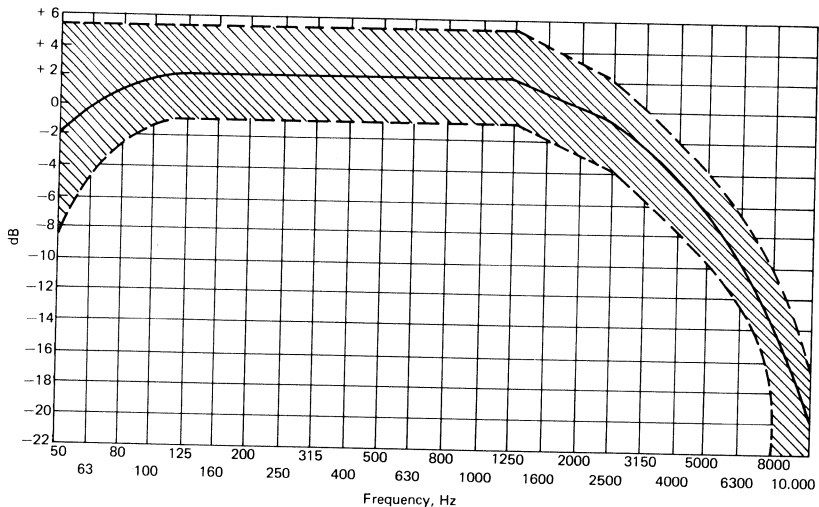
We had hoped that with the suggested minimum of labour, time and

equipment necessary for these measurements, there would be a direct appeal to the service technicians in the field and that the industry would be interested in adopting this technique as a routine check when lining up a cinema theatre. However, we did not succeed too well. Most of the work done until now, has been under the auspices of The Danish Filminstitute for as long as the free-of-charge service lasted. This service has now been terminated. The matter has been taken up also in Sweden, Norway and Finland, but it has always been a problem to secure sufficient financial support.

To carry out the measurements is a problem of a purely technical nature and it should be solved by the technicians themselves. The most obvious solution seems to be that the theatre sound equipment suppliers take over the responsibility of keeping the acoustic response characteristic within tolerance limits. Very naturally, this will not happen until an official international standard is universally accepted, hopefully in the not too distant future.

Short History of the Development of Specifications

Looking back in time, only the Motion Picture Research Council's "Standard Electrical Characteristics for Theater Sound Systems"⁹



760874

Fig. 4. Acoustic frequency response curve and tolerance limits of a monitoring chain as proposed by ISO/TC 36, WG3 (Secretariat 1) 29 (July 1969)¹⁰

(USA) tells us how to align electrically the input to the loudspeaker system. This was possible then because the choice of loudspeakers was very limited, namely, Lansing, Simplex, RCA and Western Electric. Looking at the Council's specifications we find: "the system shall not deviate by more than plus or minus 2 dB, from 50 to 8000 Hz, over the entire angle of distribution within 10 feet of the mouth of the horn." Since then, loudspeaker systems have proliferated and today not all of them keep to these specifications.

During the 8th Congress of UNIATEC (Union Internationale des Associations Techniques Cinematographiques) in Brussels, September 1968, a temporary working group was formed by some of the participating countries, including the United States, to promote further activity. An ISO meeting was held in Moscow in May 1969, where the first "Draft Proposal for Standardizing Acoustic Response of a Monitoring Chain in Motion Picture Control Rooms" was on the agenda.¹⁰ The proposed standard curve is shown in Fig.4.

Two things should be noticed here. First: the proposal is limited to motion picture control rooms only; second: it is restricted to only the "B"-chain. The term "B"-chain is adopted from an earlier suggestion by Ljungberg, which was to break the reproduction chain into two parts, A and B, as shown in Fig.5.

Ljungberg's suggestion has been universally accepted. The "A"-chain, evidently, can be lined up using ordinary sinusoidal test films, tapes or

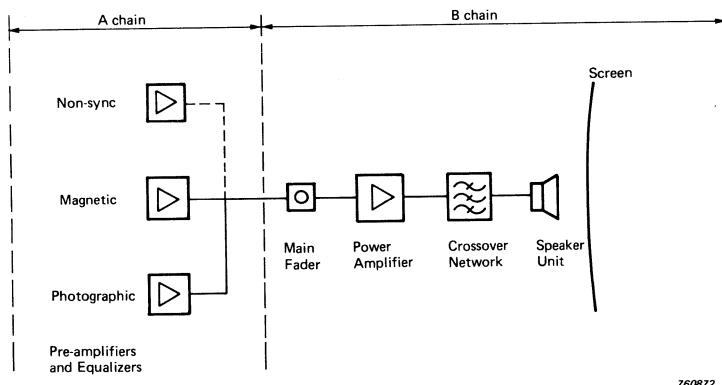


Fig.5. Division of the theatrical sound reproduction chain into an A-chain and a B-chain, as suggested by Lennart Ljungberg

discs. The now generally accepted term for the measured response of the "B"-chain alone is "acoustic response" or "electro-acoustic response". Instead, NFTU Proposal 1069 deals with the "optical-acoustic response", that is: it comprises the total reproduction chain, the sum of both the "A" and the "B"-chains. The request for an optical-acoustic standard has been strongly emphasized by the United States delegates within the ISO group. Measurement of the sound reproduction characteristic of a theatre must encompass all of the elements that can affect the character of the reproduced sound. Evidently, then, the reproduction characteristic must include the entire system from the scanning slit to the patron's ear.¹¹

During a short visit to Hollywood in April/May 1971 a number of the major studios' dubbing rooms and preview theatres were measured using the NFTU pink-noise test film. Also included was the Academy Award Theatre. The results did show a very good correlation with the NFTU proposal. Earlier discussions within ISO Working Group 3, under TC36, revealed many strong discrepancies.

Not all countries have followed the electrical roll-off suggested by the Motion-Picture Research Council. Some countries have directly refused to use it, and a third group insisted on recording and reproducing with a "flat" characteristic. This first impasse held up further progress for quite some time, until it was found that the "flat" response certainly was flat "electrically" when measured with ordinary multifrequency sine-wave test films, such as the SMPTE Type P35-MF, and when measuring across the loudspeaker dummy load. However, when measuring with a pink-noise test film and the sound level meter in the auditorium, an acoustic roll-off was found in the treble section of the loudspeaker system quite similar to the Academy roll-off. It is not an unusual situation to find the total roll-off in the loudspeaker system. We have found it in several cinemas in Denmark.

Comparison of Proposals and Results

To present, in an efficient way, a survey of all the curves proposed, and their results up to date, we present all of them here on a normalized graph with a fixed ordinate/abscissa ratio of 6 dB/octave. Only the domain above 1 kHz will be shown because this is the most discussed region. Fig. 6 shows the "Academy curve," i.e. the purely electrical characteristic of the input to the loudspeaker system.

As mentioned earlier, the NFTU proposal aims directly at an "optical-acoustic response". The aim curve above 1 kHz is shown by the dashed

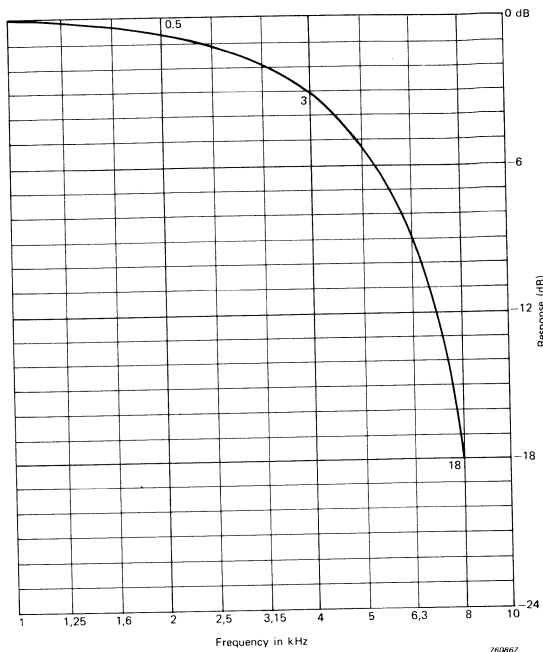


Fig. 6. The Academy response curve, representing the characteristic of the electrical input to the loudspeaker system

line in Fig. 7. For greater clarity, no tolerance limits are shown. The mean value of 100 measured public cinemas is indicated by the solid line in Fig. 7, again without tolerance limits. Only 27% of these theatres were inside tolerance limits and we must emphasize strongly that we always measured under day-to-day working conditions.¹² In many cases, high frequency response was later improved, following our suggestions, but such improvements are not included in our statistics. As can be seen in Fig. 7, the 2-kHz band is up 0,9 dB, and the 4-kHz band gave 4 dB less than the proposal. The 8-kHz band gave 3,3 dB less than the proposal.

The old prediction that optical sound would disappear from the cinema and be replaced by magnetic sound never became a reality. On the contrary, the optical soundtrack seems to hang on just as the Academy curve has stayed with us for many years. Improvements on the verge of becoming reality in optical sound reproduction seem to be

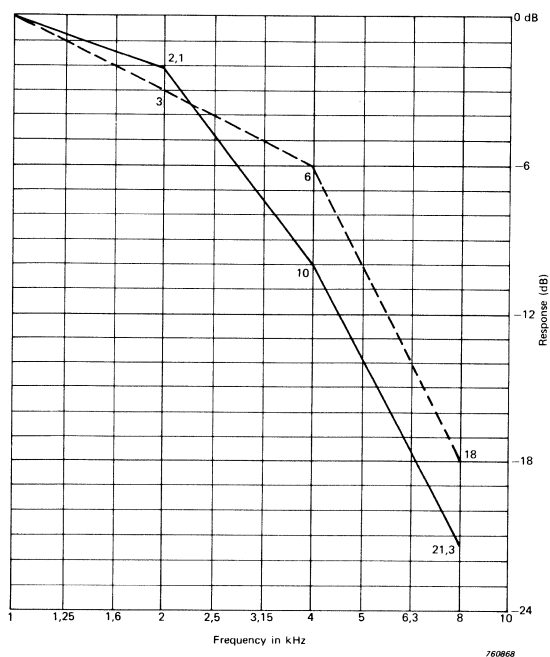


Fig. 7. Comparison of the optical-acoustic response aim curve (dashed line) of NFTU-Proposal 1069—1969 with the average response curve of 100 theatres, surveyed by NFTU (solid line)

the Dolby SVA (Stereo Variable Area) System and the new hue-modulated multichannel photographic colour soundtrack. It would be pure folly to install any of these systems without knowing the theatre's "house-curve". This curve is always ascertained and, in many cases, graphic-equalizers are inserted to improve the acoustic response.

The first ISO Draft Proposal from 1969 was never approved because it was felt that the heavy Academy roll-off should be retained. Because of current productions and in consideration of the very large number of pictures still in circulation for several years to come, the Academy roll-off could not be wiped out by the stroke of a pen. We must live with it, but on the other hand, everybody would like to take advantage of today's modern techniques.

After a long impasse in the pertinent international standardization work, new discussions took place within the ISO Preparatory Working

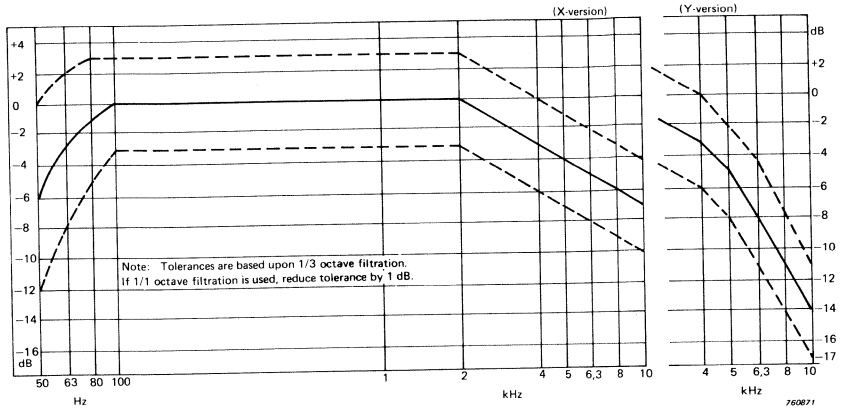


Fig.8. *New proposal for B-chain characteristic curve, agreed to by ISO preparatory working group at London meeting, 1975. (Note that this curve has a Y-version and an X-version as explained in text.)*

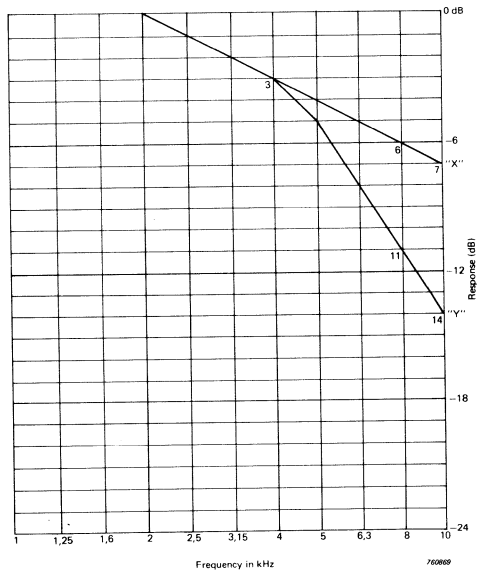


Fig.9. *Comparison of the upper range of the Y- and X-version of the 1975 ISO proposal*

Group, during its meeting in Toronto, coincident with the 116th SMPTE Conference. The groundwork was laid for a new proposal, which was later agreed to by all participants at the last meeting of the PWG in connection with Film '75 in London. In the meantime, sufficient data from theatres have become available and have been studied by the working group. Thus, the new proposal¹⁵ is not limited to monitor rooms alone, but now includes indoor theatres as well. It is not yet a complete "industry standard" because it is still restricted to cover only the "B"-chain of the total reproduction chain. No consideration has been given to the "A"-chain other than that it may be lined up using standard sine-frequency test films.

The new proposal, DIS 2969.2, is shown in Fig. 8. As can be seen, the treble section has a Y and a X version. The "Y"-version is supposed to cover pictures currently in circulation, whereas the "X"-version is intended for new sound recording systems such as the Dolby SVA or the hue-modulated colour soundtrack. Both characteristic curves are coincident below 4 kHz. Details of the upper range are more clearly shown in Fig.9. The "X"-curve should yield an excellent hi-fi reproduction in the cinema. A demonstration in the Odeon Theatre in London, during Film '75, was very convincing about all the possibilities inherent in optical sound used in the right way. The X-chain of the theatre was adjusted to the X-version of the curve of the new proposal and, as a result, dynamic range and high-frequency reproduction were most notable. One great advantage of this type of adjustment is that no high-frequency boosting on the sound negative is required to overcome "dull" cinemas. Consequently, there is a reduction in cross-modulation.

A few theatres in the Scandinavian countries have been found to have an unusual high treble section, such as the Klingenberg Kino in Oslo, or the Rialto 2 in Lyngby, Denmark, north of Copenhagen. Their treble section is shown in Fig.10. Here it should be noted, that the 1-kHz level is 2 dB low. The explanation is that the 250-Hz band is taken as the 0-dB level. As far back as 1967 Ljungberg⁸ suggested a hi-fi listening curve with an acoustic response flat from 30 to 250 Hz and from there on with a falling slope that has a gradient of $-1,5$ dB per octave; i. e., $-7,5$ dB at 8 kHz. Recently, some cinemas in Stockholm have been converted to this curve with the addition of a $65\text{-}\mu\text{s}$ roll-off. The work has been supervised by Tryggve Svensson of Stockholm.

To have roll-off or not to have roll-off has been the subject of worldwide discussion among sound technicians. It has been going on for many years and, presumably, it will stay with us for some time to

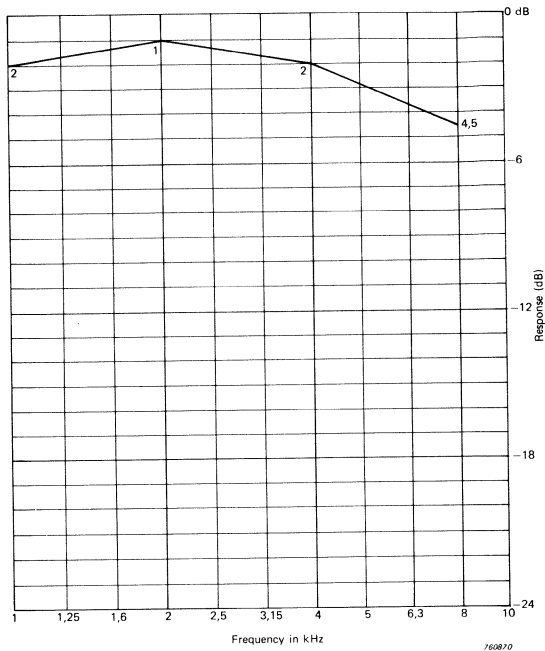


Fig.10. Treble section response curve of the 360-seat "Rialto 2" cinema in Lyngby, Denmark

come. When FM first entered radio, it was stated that a two-channel stereo up to only 7,5 kHz was preferable to a monaural channel with a range up to 15 kHz. There was something wrong with these very high and very directional frequencies, radiating from a point source. Furthermore, investigations by Henning Møller, Brüel & Kjær Co., also indicated that some roll-off was desirable.¹³

Schulein¹⁴ did research on these problems and gave an example of a subjectively determined "house-curve" for pleasing audience reaction. His response curve was quite similar to Ljungberg's⁸, only the slope starts at 1 kHz instead of at 250 Hz and the gradient is -3 dB per octave. Schulein's example relates to public address systems. He states that "at high frequencies, the human head and ears become bidirectional, with the axis of greatest sensitivity coinciding with the axis of one ear or the other. At 6,4 kHz, for example, the front-to-side sensitivity difference is typically 10 dB. As a possible aid to further measure-

ments, the construction of a microphone simulating the polar characteristic of a typical human head is worthy of further investigation”.

The electrical analogue of our present-day acoustic response measurements is a voltage reading across a complex load. The human auditory system, on the other hand, is a very complex system to consider when we want to aim for a “subjectively flat” response. To replace it with a simple (mono) objectively flat response might be to oversimplify the real conditions. Unlike home high-fidelity systems, radio and TV sound, the motion-picture industry should have an opportunity to agree on a commonly accepted replay characteristic.

The contradictory situation today seems to be that the listener has less objection to a roll-off when he is placed in the reverberant field (up to several times the critical distance) than when he is in the near field, although one should think the opposite because of distance and, consequently, screen and air losses in a cinema. We have no psycho-acoustic explanation for this, and research is still needed.

Recent Developments

Recent developments in acoustic instrumentation have made available a new one-third octave filter from Brüel & Kjær, Type 1616. It has the same mechanical dimensions as the 1613 octave filter-set, and is interchangeable with it. A portable lightweight battery- or mains-operated Graphic Level Recorder, Type 2306, weighing only 3,5 kg (7,7 lb), can now be substituted for the 2307 recorder weighing 25 kg (55 lb). Another way of lining up a theatre is to use a full bandwidth (31 Hz to 8 kHz) optical pink-noise test film (type OA 352)⁷ and then apply the filtering, either full or one-third octave, attached to the sound level meter. It will be quicker still to use a real-time analyzer. Such an instrument can be very expensive, but presumably moderately priced models will be available in the near future, designed for this type of work. A new idea for theatre loudspeaker equalization was presented by John Mosely during the 117th SMPTE Conference. It consists in the use of a limited number of variable-frequency and variable-“Q” equalizers instead of the commonly used series of 27 third-octave filters. Unfortunately, such devices are not yet on the market.

N.B. The ISO TC36 Plenary meeting in Paris, May 1976, approved the DIS 2969.2. for final letter balloting.

Acknowledgment: We would like to thank Mr. A. Tholle, Chairman of the Nordisk Film/TV Union for granting permission to reprint NFTU document 1069—1969.

References

1. BUCKLE, C. C. and A. W. LUMKIN: "The evaluation and standardization of the loudspeaker-acoustics link in motion-picture theaters", Jour. SMPTE, 78, 1058-1063, Dec. 1969.
2. NFTU-Proposal 1069—1969: Appendix 1.
3. NFTU-Proposal 1069—1969: Appendix 2, Section 2.1.
4. NFTU-Proposal 1069—1969: Appendix 2, Section 2.2.
5. NFTU-Proposal 1069—1969: Appendix 3.
6. NFTU-Proposal 1069—1969: Appendix 4.
7. Made by The Danish Filmstudio, 52 Blomstervænget, DK-2800 Lyngby, Denmark.
8. LJUNGBERG, LENNART: "Standardized sound reproduction in cinemas and control rooms", Jour. SMPTE, 78, 1046-1053, Dec. 1969.
9. "Standard Electrical Characteristics for Theatre Sound Systems", Motion Picture Research Council, Inc., Tech. Bull., 20 April 1948.
10. Subgroup ISO/TC 36/WG 3.1: "Draft proposal for standardizing acoustic response of a monitoring chain in motion-picture control rooms", Jour. SMPTE, 78, 1045, Dec. 1969.
11. VLAHOS, PETRO: "An optical/acoustic response standard for theaters and its correlation to dubbing rooms", Jour. SMPTE, 84, 3-7, Jan. 1975.

12. RASMUSSEN, ERIK: "Scandinavian efforts to standardize acoustic response in theaters and dubbing rooms", Jour. SMPTE 80, 896-899, Nov. 1971.
13. MØLLER, HENNING: "Application Notes: Relevant loudspeaker tests in studios, in hi-fi dealers' demo rooms, in homes using one-third-octave-filtered, pink-weighted random noise", Code No. 15—067, Brüel & Kjær Denmark, 1975.
14. SCHULEIN, ROBERT B.: "In situ measurement and equalization of sound reproduction systems", Jour. Audio Eng. Soc., 23, 178-186, Apr. 1975.
15. Draft International Standard DIS 2969.2.: Specifications and Measurements for Electro-Acoustic Response of Motion-Picture Control Rooms and Indoor Theatres.