

This is an electronic version of an article published in:

Heise, S. J., Verhey, J. L. & Mauermann, M., *Automatic Screening and Detection of Threshold Fine Structure*, *International Journal of Audiology* **47** (2008), 520–532, <http://dx.doi.org/10.1080/14992020802089473> .

The *International Journal of Audiology* is available online at: <http://informahealthcare.com/doi/abs/10.1080/14992020802089473>

# Automatic Screening and Detection of Threshold Fine Structure

Stephan J. Heise<sup>a</sup>, Jesko L. Verhey<sup>a</sup>, Manfred Mauermann<sup>b</sup>

<sup>a</sup>International graduate school “Neurosensory science and systems”

<sup>b</sup>Medizinische Physik  
University of Oldenburg  
Oldenburg, Germany

## Abstract

Audiograms measured with a high frequency resolution often show quasi-periodic ripples of up to 15 dB in normal-hearing listeners. This fine structure of the threshold in quiet is commonly associated with the active processes in the cochlea. Therefore its absence is discussed in the literature as an indicator of cochlear vulnerability. In order to enable a quick detection and an objective quantification of threshold fine structure, two instruments are introduced and evaluated in this article: a high-resolution tracking method for measuring fine structure (“FINESS”) and an automatic fine-structure detector (“FINESS-detector”). The method is tested on 22 subjects for its reliability, its accuracy and drifts with frequency by analysing test/retest experiments and by comparing the measured thresholds to results from a reference procedure. The results indicate that FINESS and the FINESS-detector are suitable techniques for the measurement and detection of threshold fine structure that may help to investigate further into whether fine structure is a sensitive tool for the detection of an early hearing loss.

Keywords: threshold fine structure, audiological methods

List of acronyms and abbreviations:

AFC	alternative forced choice
DPOAE	distortion product otoacoustic emission
EOAE	evoked otoacoustic emission
OAE	otoacoustic emission
SOAE	spontaneous otoacoustic emission

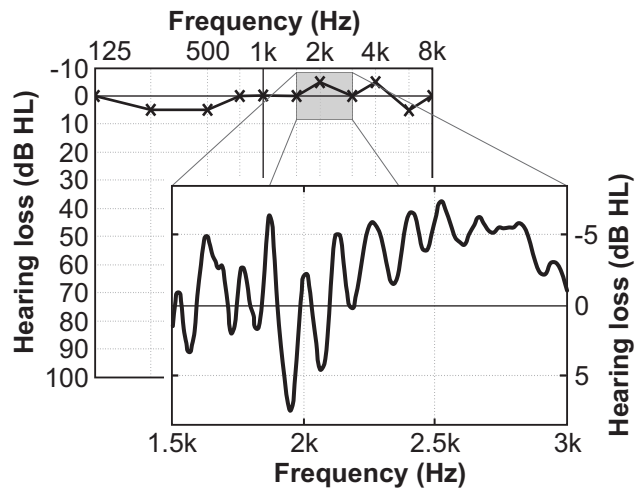
Stephan Heise  
Institut für Physik  
Universität Oldenburg  
D-26111 Oldenburg  
Email: [Stephan.Heise@uni-oldenburg.de](mailto:Stephan.Heise@uni-oldenburg.de)

## 1 Introduction

Pure tone audiograms in the standardised layout (ISO 8253-1, 1989) suggest that the threshold in quiet is flat between the audiometric frequencies. But in fact many threshold curves reveal regions with quasi-periodic threshold variations when measured with a high frequency resolution (Fig. 1). In such regions the threshold may fluctuate by as much as 15 dB over a frequency range of typically one-tenth of an octave. This phenomenon was first reported by Elliott (1958) as “ripples in the audiogram” and has since been referred to in the literature as “microstructure” (e.g. Long & Tubis, 1988b) or “fine structure” (e.g. Kapadia & Lutman, 1999; Mauermann et al., 2004). Throughout this article the term “fine structure” will be used. Effects of fine structure have also been observed in supra-threshold perception such as loudness (Mauermann et al., 2004) and modulation detection (Heise et al., 2006b). Interestingly enough, fine structure only occurs in ears — or, to be more specific, only in frequency regions — that possess normal hearing according to their audiogram. However, not all normal-hearing listeners exhibit fine structure.

The origins of threshold fine structure are assumed to be closely related to the origins of spontaneous otoacoustic emissions (SOAEs). Threshold minima, i.e. very sensitive regions of hearing, often coincide with SOAEs or large evoked otoacoustic emissions (EOAEs) (Zwicker & Schloth, 1984; Long & Tubis, 1988b; McFadden & Mishra, 1993; Horst & de Kleine, 1999). This experimental evidence for a close link between minimum spectral SOAE spacing and threshold fine structure is supported by different cochlea models which incorporate a mechanism of coherent reflection of the incoming sound at randomly distributed mechanical inhomogeneities (roughness) in the region of maximum cochlear excitation (Zweig & SHERA, 1995; Talmadge et al., 1998). Within this type of models any sound that is generated or coherently reflected in the cochlea must be conducted through the middle ear if it is to be detected as an otoacoustic emission (OAE) in the ear canal. But not all sound from the cochlea will be transmitted through the middle ear. Due to the impedance mismatch at the stapes some sound will be reflected back into the cochlea (e.g. SHERA & Zweig, 1993). The returning reflection will either enhance or partially cancel any energy at the original cochlear reflection site, depending on the round-trip travel time. If the sound was initially reflected and if it is not cancelled by the returning echo, it will be reflected again unless the properties of the cochlea have changed. Multiple internal reflections of cochlear travelling waves will occur (Zweig & SHERA, 1995). This resonance will naturally enhance the response of the basilar membrane to sounds at some frequencies, and reduce its response to sounds at others, thus resulting in threshold fine-structure minima and maxima, respectively. This same resonance behaviour can be used to generally explain the origin of the pseudo periodicity observed in all OAE fine structures with a single source, and therefore provides a common origin of OAE fine structure and threshold fine structure (see Talmadge et al., 1998).

However, a prerequisite for a significant coherent re-



**Figure 1:** Comparing threshold measurements of the same ear (subject NF-1) with different frequency resolutions. In the standard audiogram (resolution: 2 frequencies per octave) the threshold seems rather flat in the region between 1.5 and 3 kHz whereas a high-resolution measurement (100 frequencies per octave) reveals pronounced fine structure.

flection at the cochlear best sites, and therefore also for threshold fine structure, is the existence of sufficiently high and broad excitation patterns on the basilar membrane (e.g. Zweig & SHERA, 1995; Talmadge et al., 1998). At least from the model point of view the generation of broad and tall excitation patterns is sensitively related to healthy active processes in the cochlea. Therefore one may expect fine structure to be sensitive to modifications of these active processes whose progressive degeneration leads to cochlear hearing loss. Indeed several studies support this hypothesis. In patients suffering from Ménière’s disease the amount of fine structure and the degree of hearing loss have been found to be negatively correlated (Horst et al., 2003). Furst et al. (1992) observed a reduction of fine structure together with a frequency shift as a response to intense noise stimulation. Further it has been shown that aspirin consumption leads to a decrease of fine structure prior to raising the overall threshold (Long & Tubis, 1988a). Similar effects have also been found for spontaneous otoacoustic emissions (Norton et al., 1989; Furst et al., 1992; McFadden & Pasanen, 1994). All these studies indicate that fine structure may be a sensitive tool for detecting early cochlear damage, possibly even before an overall threshold shift is visible in the pure tone audiogram. However, further research is required before fine structure can be employed in clinical diagnosis.

In order to investigate the nature of fine structure, first of all a method for screening fine structure is needed. The challenge of designing such a screening technique is to keep measurement time low despite the high frequency resolution of approximately 50 frequencies per octave that is necessary for detecting fine structure. Measuring the threshold at this resolution by means of a pure tone audiogram would take around 35 minutes per octave per ear if a duration of 40 seconds per frequency is assumed. Using a psychoacoustical standard procedure (3-alternative forced choice, 3-AFC) would increase measurement time to just under seven hours per octave per ear. These durations

are impracticable for clinical use, especially when more than one octave per ear is to be screened on both ears. In the literature on fine structure therefore mostly tracking procedures have been used because of their speed.

A further prerequisite for scientific studies on fine structure is an objective means of identifying fine structure in a threshold curve. So far, in the vast majority of cases reported in the literature, regions exhibiting fine structure have simply been determined visually by the experimenter. Other studies defined a criterion for fine structure by specifying a minimum height for threshold peaks to be accepted as fine structure (e.g. Horst et al., 2003, or for DPOAE fine structure: Reuter & Hammershøi, 2006). Yet they used subjective measures to eliminate small subsidiary ripples in the threshold that do not represent ‘true’ fine structure. Therefore it is not possible to directly implement their approaches as an automatic fine-structure detection algorithm. An objective measure of fine structure would facilitate a comparison of fine-structure data between clinics and/or laboratories.

The aim of the present study was to develop an automatic method to quickly measure and quantify fine structure in an ear. Therefore an accurate estimation of the shape of a threshold curve was given priority over the acquisition of absolute threshold values. In this article we present a technique for screening fine structure which will be referred to as FINESS (*f*ine *s*tructure *s*creening). It is based on a fixed-frequency tracking procedure and contains an intelligent algorithm for controlling the repetitions of a measurement in a flexible way (Sec. 2.1). An automatic detector (“FINESS-detector”) for regions exhibiting fine structure is proposed as an objective tool for identifying fine structure in a given threshold curve (Sec. 2.2). In order to ensure a certain quality of fine-structure measurements, in general a validation of the screening technique is indispensable. Therefore the method and the detector are evaluated by a number of tests: the reproducibility of the thresholds measured with FINESS and of the results of the FINESS-detector are tested by analysing the results from test/retest measurements (Sec. 3). The accuracy of the measured thresholds is assessed by a comparison with thresholds obtained by a reference procedure (3-AFC, Sec. 2.3) in Section 4. Also by comparison with the reference procedure, the measured thresholds are checked for their stability across a wider frequency range (Sec. 5). Finally an effect of decreasing thresholds at the beginning of a measurement is closely examined (Sec. 6). The results are summarised and discussed in Section 7.

## 2 General methods

### 2.1 Screening procedure

The screening procedure FINESS described in this paper evolved from a pilot study in which several techniques for measuring fine structure were tested. A short overview of this study is included in Heise et al. (2007). The method is based on a tracking procedure, i.e. the subject is presented with a sinusoidal test tone and is instructed to press a button as long as they can hear the test tone. The level of the

tone decreases while the button is being pressed and increases while the button is in the released position. Every button press or release defines one reversal at which the test tone frequency and level are recorded. The parameter settings are given below in more detail. The FINESS application — including stimulus generation, user interface and data acquisition and analysis — is implemented as a user-friendly software programme in MATLAB.

### Parameter settings

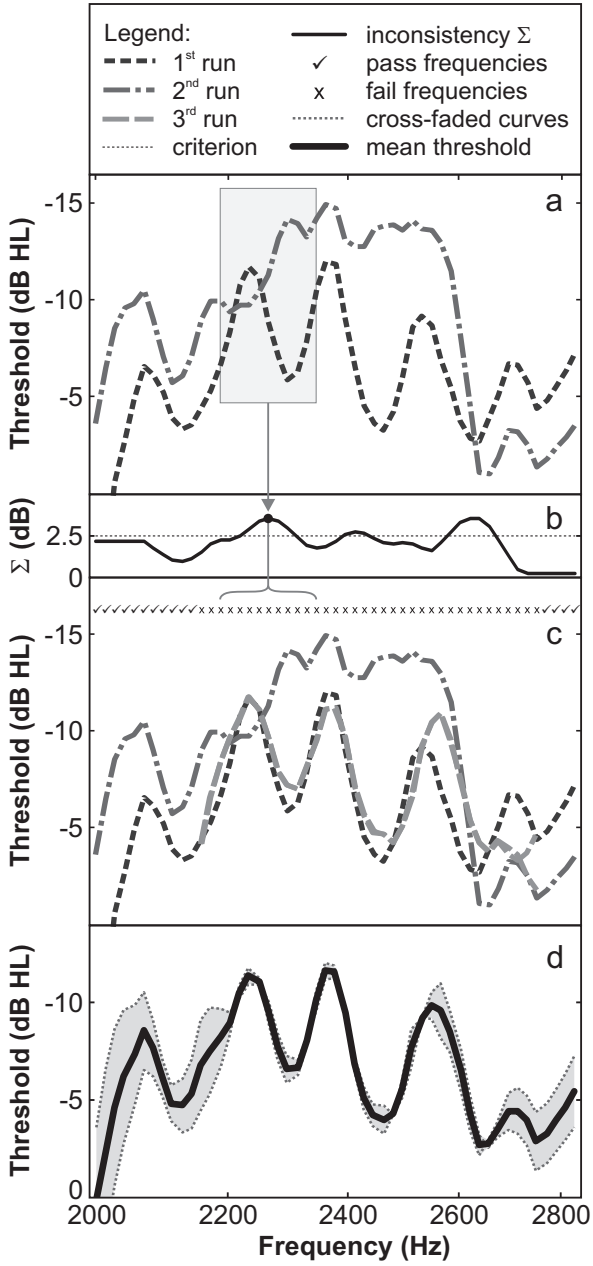
The stimuli consist of pure tones with a fixed frequency and level. The tone duration is 250 ms, which includes a 25-ms raised-cosine rise and fall (IEC 60645-1, 2001), resulting in a 200-ms steady-state stimulus in each interval. There is no silence interval between successive tones. The modulation by the cosine ramps was introduced to make the stimuli easier to detect. The frequencies are presented in ascending order. Starting from 15 dB HL the level is changed by 0.75 dB per stimulus, which is equivalent to a rate of change of 3 dB/s. This rate ensures a high level resolution while at the same time keeping the overall duration of the measurement within an acceptable limit. Besides the general interest in a short duration this is important because tracking procedures require a permanent concentration and therefore long durations may have a negative effect on the results. The direction of level change at the start of a measurement alternates between runs in order to minimise effects due to a ‘rhythmic’ response.

The frequency step size and the number of reversals recorded per frequency are two parameters that are closely related. With a small frequency step size the number of reversals per frequency may also be low because changes in the threshold from one frequency to the next are small so that the subject needs little time to adapt. In FINESS the frequency step size is set to 1/100 octave and a single reversal is recorded per frequency. That is, as in Horst et al. (2003) the frequency is changed at every reversal. A threshold estimate is calculated by smoothing the reversals, applying a locally weighted quadratic regression (“LOESS”, see e.g. Cleveland, 1979) that spans nine data points.

At the beginning of a measurement the measured thresholds tend to drop steadily before reaching a stable value. Therefore a ‘warm-up phase’ was introduced at the start of every run during which the stimulus frequency is kept constant and that merges into the actual measurement without the subject noticing. In this study the warm-up phase lasted until the changes in the measured threshold became less than 1.3 dB or until a maximum of nine reversals was reached.

### Repetitions

In order to obtain reliable threshold estimates it is necessary to repeat measurements. However, repetitions also considerably increase the overall time needed for a measurement. Therefore a strategy for repeating measurements was sought that produces reliable results in a time-efficient manner.



**Figure 2:** Illustration of the consistency check. Panel a shows the first two runs of a threshold measurement with FINNESS (subject SD-r). Obviously there are large inconsistencies in the middle of the measured frequency region. For every frequency a segment  $s$  is defined around this frequency (grey box), in which a measure  $\Sigma_s$  for the inconsistency between the thresholds in this segment is determined, resulting in the curve shown in panel b. Segments in which  $\Sigma$  exceeds a 2.5-dB criterion (dotted line in panel b) have to be remeasured (crosses in the upper part of panel c). Segments indicated by a checkmark have passed the consistency test. The remeasured curve is shown in panel c together with a copy of the curves from the first two runs from panel a. The two most consistent curves in every segment are cross-faded between the segments (dashed lines in panel d, see text for details) and averaged to give the final threshold estimate (solid line in panel d).

After the screening has been performed twice, the data is automatically checked for consistency. As it turns out the two threshold curves often agree for most of the measured frequencies. In these regions it is assumed that a reliable threshold estimate has been obtained. Only in regions that show major deviations is a third measurement taken. Note that a constant level offset between two threshold curves is accepted since the focus of the measurements lies on the shape rather than the absolute position of the threshold curves.

The consistency check is performed locally in small overlapping segments spanning 1/10 octave (i.e. 11 data points) around every measured frequency (cf. Fig. 2a). First the threshold curves of a segment are multiplied with a window in order to enhance the influence of the segment's inner frequencies. The window consists of two elements for a raised-cosine rise, a plateau of seven elements and another two elements for a raised-cosine fall (the same window is also used later for cross-fading the data; see next paragraph). Next, as a measure of the consistency of two threshold curves, the variation of the distance between these curves within a segment is analysed by calculating the standard deviation of the difference of the normalized data:

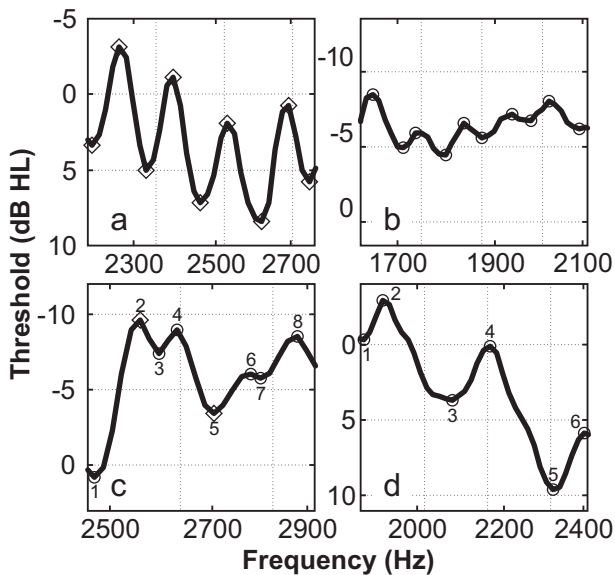
$$\Sigma_s := \sum_{i=1}^n \frac{(\Delta_i - \frac{1}{n} \sum_{k=1}^n \Delta_i)^2}{n-1}$$

with

$$\Delta_i := \left( w_i T_1(f_i) - \frac{1}{n} \sum_{k=1}^n w_k T_1(f_k) \right) - \left( w_i T_2(f_i) - \frac{1}{n} \sum_{k=1}^n w_k T_2(f_k) \right)$$

where  $T_1(f_i)$  and  $T_2(f_i)$  denote the two thresholds at the frequency  $f_i$  in the segment  $s$ ,  $n = 11$  is the number of data points in a segment and  $w_i$  denote the elements of the window function. If  $\Sigma_s$  exceeds a critical value of 2.5 dB in a segment — i.e. if the distance between the thresholds fluctuates too much — the measurements are considered inconsistent in this segment (Fig. 2b). If an adjacent segment also shows inconsistent thresholds, the frequencies in the current segment are marked for future remeasuring. After all segments have been analysed, a third run is performed in which the threshold is remeasured at the frequencies of those segments that did not pass the consistency check (Fig. 2c).

For the final threshold estimate first the two most consistent measurements are determined in each segment (i.e. the two thresholds that minimise  $\Sigma_s$ ). These will be referred to as  $T_\alpha(s)$  and  $T_\beta(s)$  in the following. In order to achieve a smooth transition between the thresholds in overlapping segments, the thresholds are multiplied with the window function described above and cross-faded in the following way: for each frequency the thresholds  $T_\alpha(s)$  are averaged across the segments containing this frequency, resulting in a curve  $T_\alpha$ . The same is done for  $T_\beta(s)$ . This is illustrated in Figure 2d where the two curves  $T_\alpha$  and  $T_\beta$  are shown as dotted lines. When comparing them with the three curves in Figure 2c, it can be



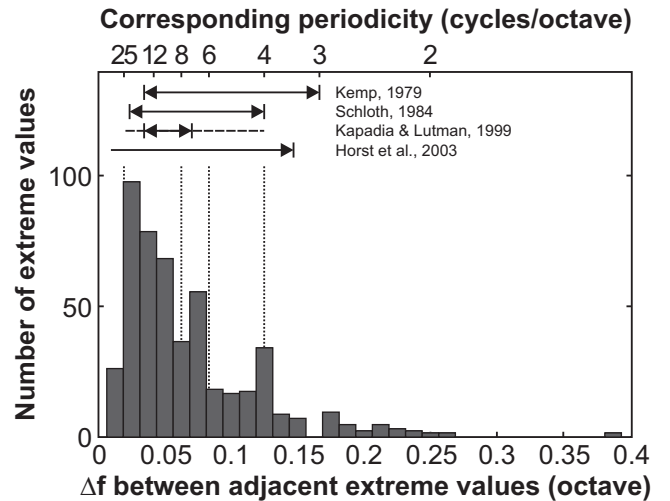
**Figure 3:** Four segments of threshold curves that illustrate how the basic algorithm of the fine-structure detector works (the thresholds are from subjects ZM-l, MP-l, FH-r and MJ-l). The extreme values in the threshold curves which are analysed by the detector are indicated by diamonds (extreme values that meet the detector’s criterion) and circles (others). The extreme values are numbered in panels c) and d) for reference purposes (see text). When applying a 3-dB/10-octave criterion (grid), threshold curves a) and c) are classified as fine structure while curves b) and d) are not.

seen that at the beginning  $T_\alpha$  and  $T_\beta$  equal the threshold curves from the first two runs. Then around 2200 Hz they blend over to represent the threshold curves from runs one and three, until they blend again at 2600 Hz and become the threshold curves from runs one and two again. In a last step  $T_\alpha$  and  $T_\beta$  are averaged to give the final threshold estimate. As an estimate of the measurement uncertainty,  $T_\alpha$  and  $T_\beta$  are given in the graphs as boundaries of a shaded area (see Fig. 2d).

## 2.2 Fine-structure detector

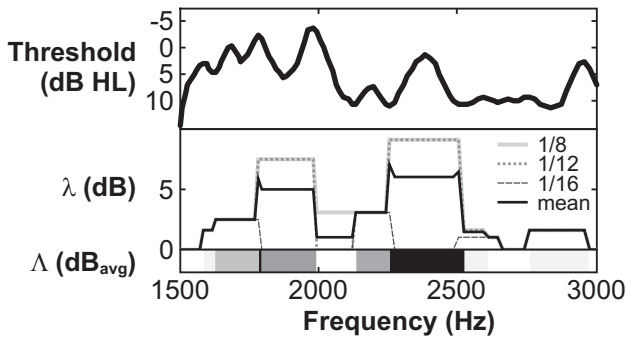
Whether or not and where a threshold has fine structure is typically judged by visual inspection. Since this measure of fine structure is rather subjective, it is somewhat difficult to compare fine-structure data gathered by different research groups. In order to provide a more objective method for the classification of fine structure, an algorithm that automatically identifies regions exhibiting fine structure in a given threshold curve is developed in the following. This algorithm, referred to as FINNESS-detector, is implemented as a separate module so that it can be applied to any threshold curve.

In a first step, the detector analyses the extreme values of a threshold curve with respect to their frequency and level spacing. Three parameters need to be specified: a minimum level difference  $\Delta L_{\min}$  and a range of allowed frequency differences defined by  $\Delta f_{\min}$  and  $\Delta f_{\max}$ . If a peak is more than  $\Delta L_{\min}$  and between  $\Delta f_{\min}$  and  $\Delta f_{\max}$  apart from its neighbouring troughs, then the threshold is classified as displaying fine structure in the frequency



**Figure 4:** Frequency difference distribution of adjacent extreme values. 487 extreme values from 36 threshold curves (each covering one octave) were analysed. A frequency difference of  $\Delta f$  between adjacent extreme values corresponds roughly to a fine-structure periodicity of  $(2 \cdot \Delta f)^{-1}$  which defines the scale of the top abscissa. A bin size of 1/80 octave was used for the histogram. The critical values for  $\Delta f_{\min}$  and  $\Delta f_{\max}$  that are used in the FINNESS-detector are indicated as vertical dotted lines. Further the periodicity regions found in the literature are given (see text for more details).

region covering these three extreme values. This is illustrated in Figure 3 for  $\Delta L_{\min} = 3$  dB,  $\Delta f_{\min} = 1/50$  octave and  $\Delta f_{\max} = 1/10$  octave. Extreme values that meet the criterion are depicted as diamonds, others as circles. In Figure 3a every extreme value is more than 3 dB and between 1/50 and 1/10 octave from its nearest neighbours. Therefore the entire range is classified as fine structure. The threshold curve in Figure 3b also shows ripples but in this case they are not deep enough to count as fine structure when applying a 3-dB criterion. Figure 3c displays a threshold segment that contains pairs of extreme values which should be ignored when searching for fine structure (e.g. nos. 6 & 7). Such extreme values will be referred to as ‘subsidiary’ extreme values in the following. The algorithm accounts for this by analysing not only distances between nearest neighbour extreme values but also between next-nearest neighbour, next-next-nearest neighbour, etc. extreme values. For example, when analysing extreme value no. 5 in Figure 3c, the level difference between nos. 5 and 6 is found to be less than the chosen criterion of 3 dB. Hence the level difference between nos. 5 and 8 is examined which is more than 3 dB. Because the level difference between nos. 6 and 7 is less than 3 dB, extreme values nos. 6 and 7 are labelled as subsidiary extreme values and excluded from further analysis. The same holds for extreme values nos. 3 and 4. Note that the definition of a subsidiary extreme value depends on  $\Delta L_{\min}$ . Since the remaining extreme values (nos. 1, 2, 5 and 8) meet the criterion, this threshold segment is classified as fine structure by the detector. Figure 3d finally shows a situation in which extreme values nos. 2 and 3, and nos. 4 and 5 are more than 1/10 octave apart so that no fine structure is detected.



**Figure 5:** Illustration of the fine-structure measure  $\lambda$ . A threshold curve from one subject (HR-r) is drawn in the top panel. The bottom panel shows a measure  $\lambda$  of the fine structure in that threshold curve for four different periodicity criteria:  $\Delta f_{\max} = 1/8, 1/12$  and  $1/16$  octave and the arithmetic mean  $\Lambda$  of these three. Below this panel the average measure  $\Lambda$  is shown again coded as grey-scales, where white means less than  $1 \text{ dB}_{\text{avg}}$  and black more than  $6 \text{ dB}_{\text{avg}}$ .

The basic algorithm just described returns frequency regions containing fine structure according to a fixed criterion. However, such a fixed criterion is not ideal for describing fine structure. For instance, this could lead to a situation in which a 3.0-dB peak is classified as fine structure while a 2.9-dB peak is not. Or a peak–trough distance of  $1/10$  octave may be accepted while a distance that is larger by only 1 Hz is not. Besides, the definition of fine structure in the literature is not at all clear cut. For example, a variety of values has been proposed for the critical peak depth in the few studies that dealt with the problem of quantifying fine structure: Kapadia & Lutman (1999) interpreted ripples deeper than 2 dB as fine structure, Horst & de Kleine (1999) used  $\Delta L_{\min} = 5$  dB, whereas Horst et al. (2003) applied a 3-dB criterion. Therefore in a second step a more gradual characterisation of fine structure is developed.

First the basic algorithm is run several times on the same threshold curve for a series of values for  $\Delta L_{\min}$  from 0 through 10 dB in steps of 0.5 dB. In this way a critical value  $\lambda$  for  $\Delta L_{\min}$  can be found for every measured frequency so that the region around that frequency is classified as fine structure when the algorithm is run with  $\Delta L_{\min} \leq \lambda$ , but not when  $\Delta L_{\min} > \lambda$ . For example, frequencies around a 5-dB peak will have  $\lambda = 5$  dB. These critical values  $\lambda(f)$  may be interpreted as a local measure of fine structure for a given  $\Delta f_{\min}$  and  $\Delta f_{\max}$ .

However, the range of accepted frequency differences should also be less rigid. The periodicity range found for fine structure in the literature varies from study to study. An overview of these ranges is given in the upper part of Figure 4. Kemp (1979) reported a spacing of (loudness) maxima between  $1/3$  and  $1/14$  octave. Schloth (1983) found values between  $1/4$  and  $1/21$  octave for the spacing of threshold minima. Kapadia & Lutman (1999) performed a Fourier transform of the log-scaled threshold curve which showed a broad peak between 7 and 14 cycles/octave. In a further analysis they examined the ripple depth in bandpass-filtered threshold curves. The pass band of their filter — i.e. the periodicity range they

accepted for fine structure — was between 4 and 24 cycles/octave (dashed line in Fig. 4). Similarly Horst et al. (2003) obtained a ‘global’ threshold by applying a moving-average smoothing and subtracted this ‘global’ threshold from the threshold curves prior to analysing the ripple depths. This is equivalent to highpass-filtering the threshold curves. The size of their smoothing window corresponds to a cut-off frequency of 3.4 cycles/octave. For the development of the FINESS-detector we analysed the periodicity of fine structure by examining the frequency difference of adjacent extreme values in 36 threshold curves that had been measured with FINESS over one octave (20 of these curves had been measured for experiment III in Sec. 5, and 16 had been obtained during a different study). Only extreme values that were not identified as subsidiary peaks when running the basic algorithm with  $\Delta L_{\min} = 3$  dB were considered. The distribution of the frequency differences is shown in Figure 4. The upper abscissa gives the corresponding periodicity — i.e. the inverse of the double frequency difference — in cycles per octave. The most frequent periodicities are around 17 cycles/octave and the distribution gradually decays towards lower periodicities. The data agree with the periodicity ranges found in the literature (as indicated by the arrows in Fig. 4). In order to account for this distribution in the FINESS-detector, the fine-structure measure  $\lambda(f)$  developed above is calculated three times for three different  $\Delta f_{\max}$  ( $1/16, 1/12$  and  $1/8$  octave) and a constant  $\Delta f_{\min}$  ( $1/50$  octave, cf. dotted lines in Fig. 4), and then averaged. This process is illustrated in Figure 5. The averaged measure is denoted by  $\Lambda(f)$ .  $\Delta f_{\min}$  was set to  $1/50$  octave since periodicities above 25 cycles/octave are considered as ‘noise’. Thus periodicities between 25 and 8 cycles/octave are weighted 1.5 times as much as periodicities between 8 and 6 cycles/octave which again receive twice as much weight as periodicities between 6 and 4 cycles/octave. Note that by incorporating multiple accepted periodicity ranges, the interpretation of the units of the fine-structure measure has become less straight forward. A value of, say,  $\Lambda = 4$  only implies a 4-dB peak if the peak was rated as 4-dB peak for all three periodicity ranges. However, it could also result from a 6-dB peak that is too broad to count as fine structure when applying  $\Delta f_{\max} = 1/16$  octave (i.e. as an average of 6, 6 and 0 dB). Therefore the units of  $\Lambda$  are denoted as “ $\text{dB}_{\text{avg}}$ ” in this article.

The local measure of fine structure obtained in this way is represented as a grey-scaled bar (see e.g. Fig. 6) where black indicates pronounced fine structure of more than  $6 \text{ dB}_{\text{avg}}$ , and white indicates regions without fine structure, i.e. with fine structure of less than  $1 \text{ dB}_{\text{avg}}$ . If a fixed fine-structure criterion is desired, regions in which the fine-structure measure  $\Lambda$  exceeds a critical value of  $3 \text{ dB}_{\text{avg}}$  may be highlighted or indicated as black bars like in the upper part of each panel in Figure 6. In contrast to the fine-structure regions returned by the basic algorithm, these regions result from a combination of criteria for the accepted periodicity range, thus accounting for the distribution of fine-structure periodicities (Fig. 4).

### 2.3 Reference procedure

For judging the quality of the thresholds measured with FINESS an adaptive three-alternative forced choice (3-AFC) technique was used as a reference procedure. This method is known to produce thresholds with minimal bias and is well established in psychoacoustics. Due to its long measurement time, however, it is unsuitable for the screening of fine structure (see Introduction).

The subject is presented with three intervals that are visually highlighted, one of which contains the stimulus. The subject has to indicate in which interval they heard the stimulus. After two correct answers the stimulus level is decreased, whereas a wrong answer leads to an increase of stimulus level. This is known as the 1-up 2-down paradigm which is the most efficient paradigm in 3-AFC procedures (Kollmeier et al., 1988). In this way the stimulus level required for 70.7% correct detection is estimated (Levitt, 1971).

The stimuli used for the reference measurements equalled those used for FINESS, i.e. they were pure tones lasting 250 ms, including 25-ms raised-cosine ramps. The inter-stimulus interval was set to 500 ms. Following an answer visual feedback was given. The stimulus level started at 15 dB HL and changed in steps of 6 dB. After the second reversal the step size was reduced to 3 dB and at the beginning of the measurement phase after the fourth reversal to 1 dB. In the measurement phase eight reversals were recorded which were averaged to produce a threshold estimate. To ensure a certain quality of these estimates each run was repeated several times until the last two (Sec. 4) or three (Sec. 5) estimates for each frequency had a standard deviation of less than 4 dB.

### 2.4 Subjects and experimental set-up

In total 22 subjects (14 male, 8 female) aged between 22 and 45 took part in this study (see Table 1 for details). The subjects had no reported hearing impairments. In the frequency range covered by the experiments in this study (1–3 kHz) all subjects showed thresholds of 15 dB HL or better in their pure tone audiograms.<sup>1</sup> Different subsets of subjects participated in the four experiments: all 22 subjects participated in the first experiment, 15 (8 male, 7 female) in the second experiment, 20 (13 male, 7 female) in the third experiment and 20 (14 male, 6 female)

<sup>1</sup>Subjects with thresholds in quiet of up to 15 dB HL according to their audiogram are generally considered to have normal hearing. At 8 kHz three subjects (HJ, HS, NM) showed thresholds of 20 dB HL and one subject (HR) of 25 dB HL in one ear. This may be interpreted as a mild hearing loss at this high frequency, although some studies even consider a threshold of 25 dB HL as still normal. In addition, it is possible that some of the subjects may already have an initial hearing loss in the relevant frequency range of 1–3 kHz which is undetectable by standard audiometry. The aim of the present study is to develop a reliable method for measuring fine structure in order to help to possibly establish a link between such early hearing losses and the amount of threshold fine structure in future studies. Thus, for evaluating the method it is not crucial that all subjects have normal hearing, as long as some of them show fine structure and some do not. In order to enhance the probability of finding threshold fine structure, only subjects with no (or only a mild) hearing loss in their pure tone audiograms were chosen for the present study.

label	sex	age [y]	ear	threshold at ... kHz [dB HL]							
				.125	.25	.5	1	2	4	8	
FE	f	23	l	5	0	0	0	0	0	0	5
			r	0	0	0	0	5	0	5	
FH	f	23	l	0	0	0	5	5	10	5	
			r	0	0	-10	0	-5	0	5	
HA	f	23	l	10	5	5	0	0	5	5	
			r	5	10	10	5	10	10	15	
HJ	f	30	l	0	0	0	-5	0	-5	20	
			r	0	-5	-5	0	5	5	10	
HM	m	30	l	0	-5	-5	-5	0	0	0	
			r	0	-5	-5	-5	0	0	10	
HR	m	33	l	-5	0	0	0	10	5	20	
			r	0	-5	0	0	5	10	25	
HS	m	29	l	10	0	0	-5	0	-10	15	
			r	5	0	0	-5	-5	-5	20	
JM	m	45	l	5	5	10	0	5	5	10	
			r	0	0	0	0	0	10	10	
KS	m	23	l	0	0	0	-5	0	-5	5	
			r	0	0	-5	-5	0	-5	15	
LE	m	24	l	5	0	5	0	5	5	0	
			r	0	-5	0	0	5	5	0	
MA	m	24	l	0	-5	0	0	0	-10	5	
			r	0	0	0	0	5	-5	10	
MJ	f	25	l	0	0	0	0	0	5	5	
			r	10	10	10	10	10	15	10	
MP	m	24	l	-5	-5	-10	-5	-5	5	5	
			r	-5	-5	0	0	-5	5	5	
NF	m	26	l	0	5	5	0	-5	-5	0	
			r	5	0	0	0	5	-5	0	
NG	m	40	l	0	0	5	0	0	0	5	
			r	0	-5	0	0	5	0	15	
NM	m	32	l	-5	-5	-5	-5	-5	5	20	
			r	-5	0	-5	0	5	5	0	
SD	m	25	l	0	0	0	-5	0	5	0	
			r	0	-5	-5	-5	5	0	0	
TA	f	23	l	5	15	15	5	-5	5	5	
			r	0	0	5	-5	5	-5	0	
WA	f	22	l	0	5	5	-5	0	0	5	
			r	5	0	0	0	0	0	5	
WJ	m	24	l	10	5	10	5	5	5	15	
			r	10	5	10	0	10	0	10	
WS	m	25	l	5	0	0	5	5	5	5	
			r	10	0	-5	5	0	5	5	
ZM	f	23	l	5	0	0	5	5	5	10	
			r	0	-5	-5	0	5	-10	5	

**Table 1:** List of subjects. For each subject the acronym, gender (f=female, m=male), age and standard pure tone audiogram data for left (l) and right (r) ear are given.

in the fourth experiment. The subjects were seated in a sound-treated booth. The signals were generated digitally in MATLAB, amplified by a Tucker Davis HB7 amplifier and played back via Sennheiser HDA200 headphones. The headphones were calibrated on a Brüel & Kjær artificial ear (type 4153) and corrected for the reference equivalent threshold sound pressure level (ISO 389-8, 2004).

### 3 Experiment I: Reliability

In order to verify the reliability of the data obtained by FINESS, test/retest measurements were carried out. The aim was to check how reproducible the thresholds and in particular their fine structures are. For this purpose thresholds were measured twice with FINESS over a frequency range of half an octave starting at 2 kHz. The duration of the experiment was kept short by measuring only half an octave and by taking the second measurement immediately after the first so as to minimize the influence of potential fluctuations in the thresholds over time.

The mean duration for the measurement of half an octave was 5 min ( $\pm 1$  min). The individual results of the test/retest measurements are shown in Figure 6 for all 22 subjects who participated in the experiment. In general the threshold curves show a good reproducibility in terms of the overall shape and the position and depth of fine-structure ripples. This is confirmed by the correlation coefficient  $r$  of the two threshold curves of every subject (see Fig. 6) which rates the similarity of the threshold curves:

$$r(T_1, T_2) = \frac{(T_1 - \bar{T}_1) \bullet (T_2 - \bar{T}_2)}{|T_1 - \bar{T}_1| \cdot |T_2 - \bar{T}_2|}$$

where  $T_1$  and  $T_2$  are vectors containing the values of the two threshold curves with mean values  $\bar{T}_1$  and  $\bar{T}_2$ , and  $\bullet$  represents the dot product of two vectors. Sixteen subjects show correlations above 0.85 and only two show correlations below 0.7. The average correlation coefficient is 0.88.<sup>2</sup>

Since FINESS is designed to measure fine structure it is important to know how reliably the position and the depth of fine-structure peaks can be reproduced. To address this question the extreme values in the threshold curves were analysed with respect to their frequency and their depth which was defined as the mean level difference to the adjacent extreme values. For the following analysis subsidiary extreme values as defined in Section 2.2 for a 3-dB criterion were ignored. In total 249 pairs of extreme values were examined. On average the frequencies of corresponding extreme values in the test and the retest measurements deviate by 0.03% (or 1/2000 octave) with a standard deviation of 0.90% (or 1/77 octave). This is similar in size to the frequency step size in FINESS, which

<sup>2</sup>Correlation coefficients in this study were averaged by applying Fisher's transformation  $Z = \text{artanh}(r)$ , then calculating the arithmetic mean of the transformed coefficients and transforming it back. This is done because coefficients of correlation are not values on a scale of equal metric units. Differences between large  $r$ 's are actually much greater than those between small  $r$ 's. Therefore the sampling distribution of the correlation coefficient is skewed so that the arithmetic mean of correlations would underestimate their central tendency (?Silver & Dunlap, 1987).

is 1/100 octave, and therefore negligible. The depth of the extreme values deviates by  $-0.1 \pm 1.8$  dB (mean and standard deviation) between the test and the retest measurements. These occasional deviations are distributed approximately normally which means that deviations of more than 3 dB occur in only 9% of the cases.

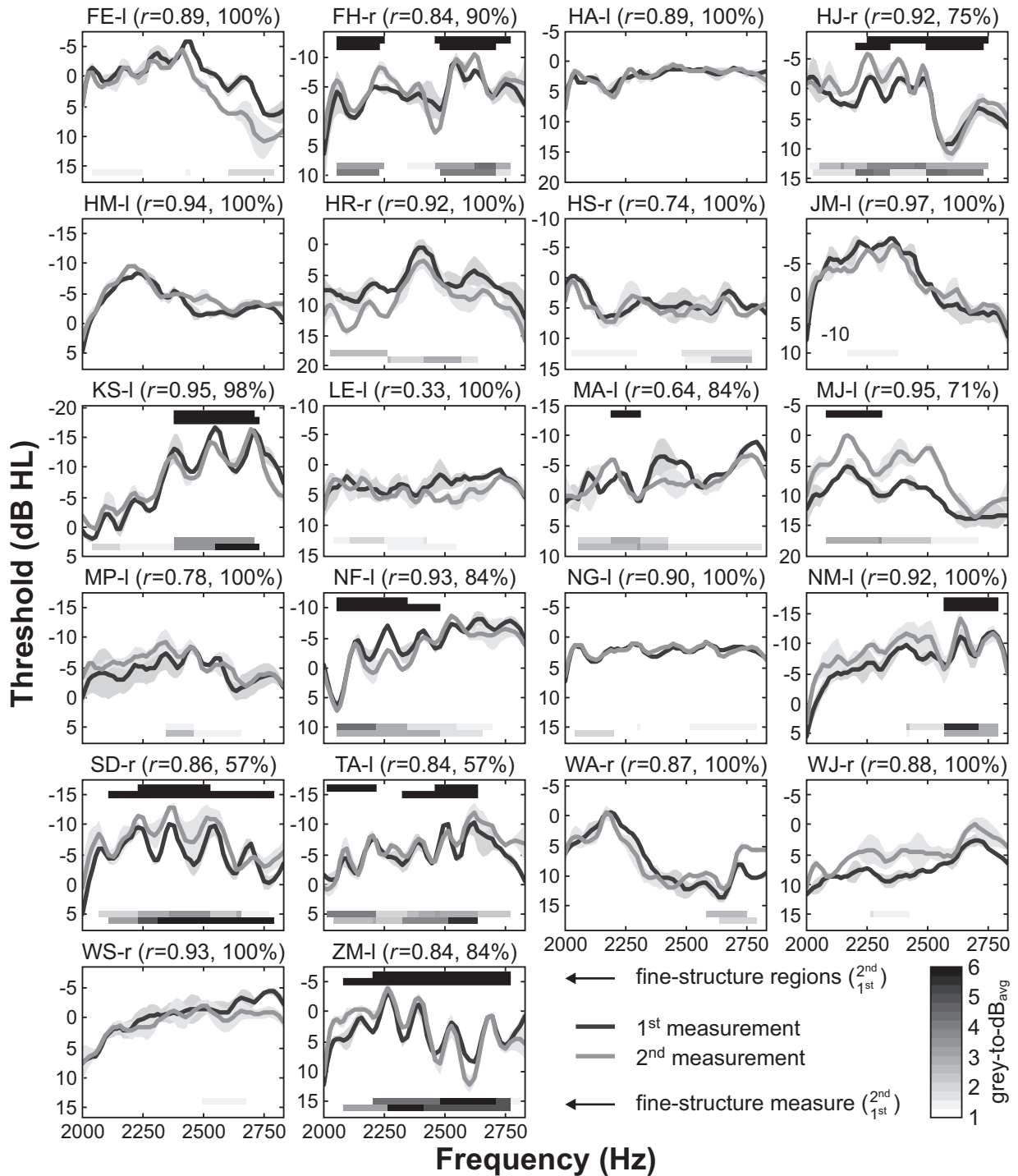
In order to test whether these occasional deviations affect the reproducibility of the results of the FINESS-detector, fine-structure measures as well as fine-structure regions are shown in Figure 6. The fine-structure measures are displayed as grey-scales in bars beneath each pair of threshold curves, black standing for very pronounced fine structure (more than 6 dB<sub>avg</sub>) and white for very weak fine structure (less than 1 dB<sub>avg</sub>). Regions in which this measure exceeds 3 dB<sub>avg</sub> were marked as fine-structure regions by a black bar above the threshold curves. In both cases the lower bar represents the first measurement and the upper bar the second. First of all it should be noted that the FINESS-detector detects clear fine structure when it is present as e.g. in the subjects KS-l, SD-r or ZM-l. In the opposite case of clear absence of fine structure — e.g. subjects HA-l, HM-l or WS-r — no fine structure is detected. And different strengths of fine structure are represented as different shades of grey in the fine-structure measure which can be verified best in subjects KS-l and NF-l. When comparing test and retest measurements, the similarity between corresponding fine-structure measures is evident in most cases. However, small differences in the detected strength of the fine structure do exist and are probably due to the small fluctuations in fine-structure depth mentioned above. The identified fine-structure regions usually overlap to a high degree. The percentage of equally judged frequency regions is given for every subject in Figure 6. The good agreement of the results is reflected in a high mean correspondence of 91% ( $\pm 14\%$ ). All in all the FINESS-detector seems robust against small fluctuations in the measured thresholds and reliably detects fine structure.

### 4 Experiment II: Accuracy

The accuracy of the thresholds obtained by FINESS was tested by comparing them to results from a 3-AFC reference procedure (see Sec. 2.3). This comparison focussed on the shape of the threshold curves, since tracking procedures such as FINESS are not expected to yield precise estimates of the absolute threshold due to the variability of the subjects' internal threshold criteria (Marshall & Jesteadt, 1986). If, for instance, a subject only presses the response button when they are absolutely sure that they clearly perceive the test tone, the measured threshold values will lie slightly above the absolute threshold. However, up to about 20 phon this should hardly affect the shape of the threshold curves except for a slight smoothing (Mauermann et al., 2004).

The comparison was made in a rather small frequency range of about one-fifth of an octave. For each subject the threshold was first measured with FINESS. Then two threshold minima, i.e. very sensitive regions, were chosen, around which threshold estimates were obtained at





**Figure 6:** Individual test/retest results from 22 subjects. Together with the two threshold curves (lines) and their errors (shaded areas, cf. Fig. 2d) the results from the FINESS-detector are shown. Fine-structure measures are represented as grey-scales in bars below the threshold curves, whereas the fine-structure regions (when applying a 3-dB<sub>avg</sub> criterion) are displayed as black bars above the threshold curves. In both cases the lower bar corresponds to the first measurement and the upper bar to the second. For every subject the correlation coefficient  $r$  of the two threshold curves and the overlap between the detected fine-structure regions as a percentage of the whole frequency region are given in the panel title.

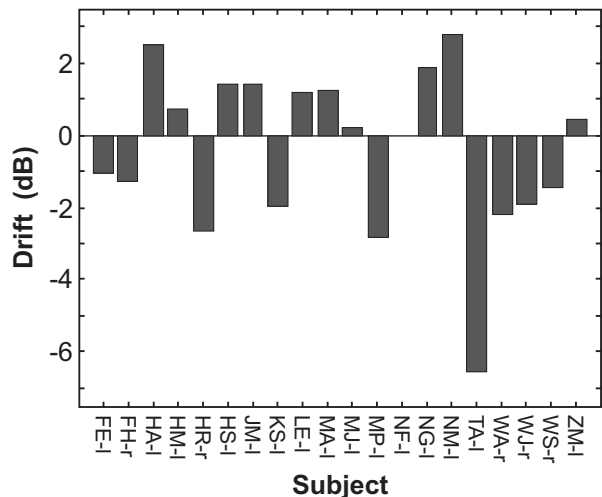
nine frequencies by a 3-AFC procedure. The frequencies were equally spaced on a linear scale and placed so that the second and the eighth frequency corresponded to the two minima. When no fine structure was present the frequencies were chosen in a way that they covered a range of about one-fifth of an octave. Since the frequencies' positions were defined by the individual thresholds, their spacing was different for each subject. For each frequency two threshold estimates were obtained by the 3-AFC procedure.

The individual results for all 15 subjects are shown in Figure 7. In order to facilitate the comparison, the thresholds measured with FINESS (thin lines) were shifted vertically (thick lines) so as to minimize the sum of their squared distance to the 3-AFC data (circles). The shapes of the threshold curves measured with FINESS and the 3-AFC procedure are very similar. As a measure of the strength of the relationship between the data from the two methods, correlation coefficients were calculated for each subject. These are given in Figure 7 in the bottom left corner of each panel. In four subjects the correlation between the thresholds obtained by FINESS and the 3-AFC procedure is higher than 0.9, and only in two subjects it is less than 0.7. The average correlation<sup>2</sup> is 0.87.

## 5 Experiment III: Stability across frequencies

Since the frequencies are measured ascendingly in FINESS, gradual changes in the thresholds with time will lead to drifts in the thresholds across frequencies. Such drifts could result for instance from fatigue or adaptation. In order to investigate whether such drifts are present in FINESS, thresholds were measured over a larger range than in Section 4, namely one octave, and compared to thresholds from a 3-AFC procedure (see Sec. 2.3). Two frequencies were chosen for the 3-AFC measurements, one in the lower and one in the upper quarter of the octave:  $f_1$  and  $f_2$ . Whenever possible they were placed at positions where the FINESS threshold showed a high consistency (i.e. a small error shade). For each of the two frequencies three threshold estimates were collected with the 3-AFC method and averaged.

On average the 20 subjects took 10 min ( $\pm 2$  min) for the FINESS measurement and 17 min ( $\pm 4$  min) for the 3-AFC measurement. As in Section 4 the thresholds measured with FINESS were shifted vertically in order to minimize the sum of their squared difference to the 3-AFC data. If global trends should exist then the shifted FINESS thresholds should either always lie above those measured with the 3-AFC procedure at  $f_1$  — and thus below them at  $f_2$  — or vice versa. Therefore the thresholds may be analysed for drifts by examining the difference between the FINESS and the 3-AFC data at one of the two frequencies. This difference is plotted in Figure 8 for  $f_2$ . For 10 out of 20 subjects the difference is positive, and for the other 10 it is negative. Excluding subject TA-1 who shows an exceptionally large drift of  $-6.6$  dB, all values lie within  $\pm 3$  dB. On average (again excluding TA-1) the FINESS thresholds are 0.1 dB below the 3-AFC



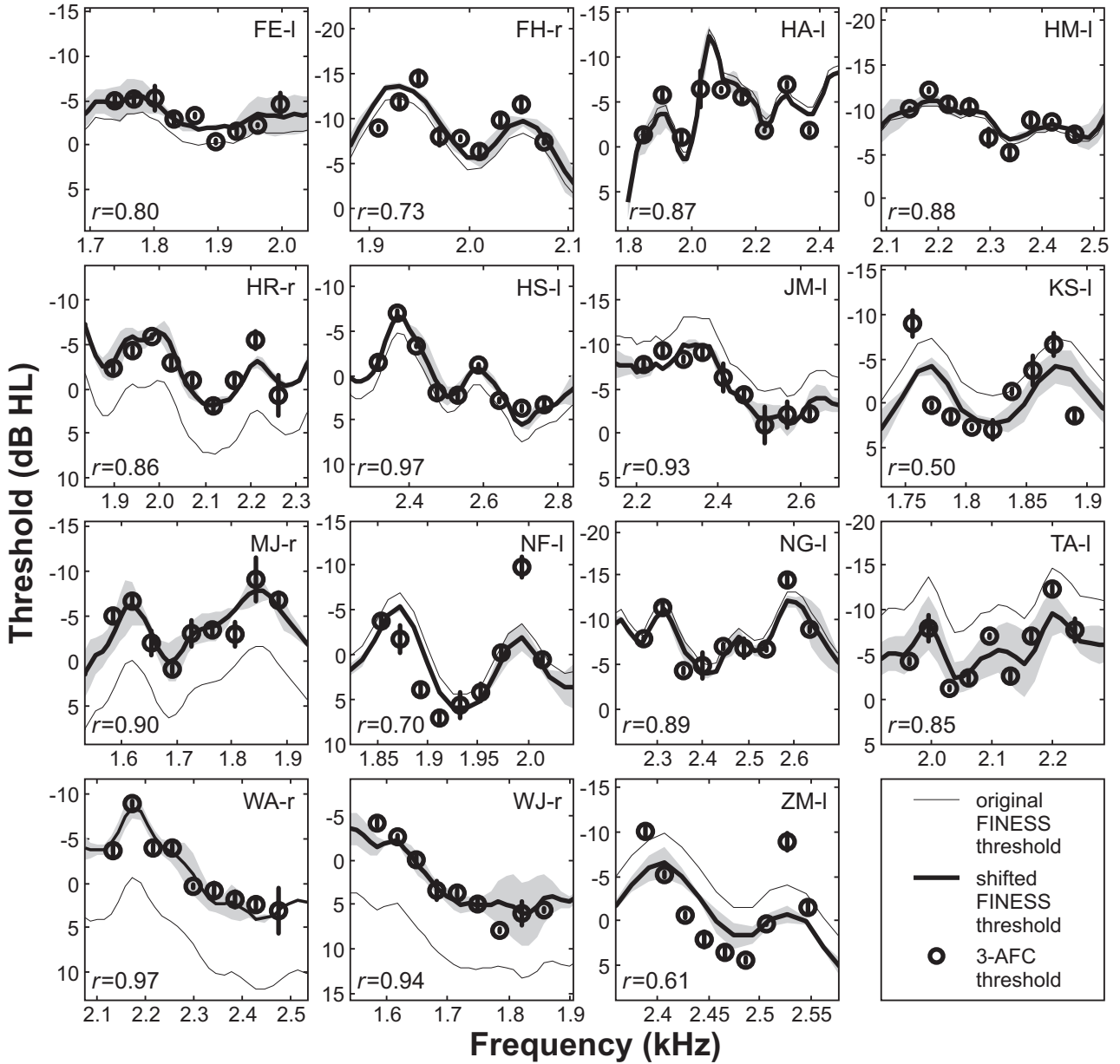
**Figure 8:** Testing for drifts across frequencies. The thresholds measured with FINESS and the 3-AFC procedure were compared at two widely separated frequencies. The difference between the corrected FINESS threshold and the 3-AFC threshold at the second frequency is plotted for each subject.

thresholds at  $f_2$ , which corresponds to a mean drift of  $-0.1$  dB/octave ( $\pm 2.5$  dB/octave). That is, the FINESS threshold estimates tend to decrease slightly over time.

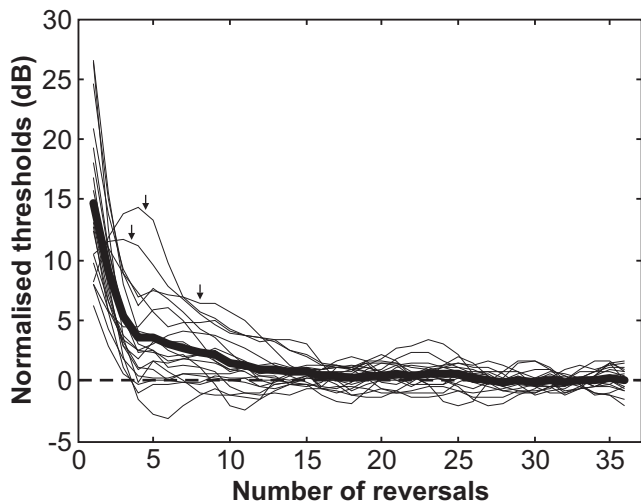
## 6 Experiment IV: Pre-measurement phase

As was mentioned in Section 2.1 the subjects are given some time at the beginning of each FINESS run to get accustomed to the procedure. This warm-up phase lasted until the measured threshold changed less than 1.3 dB or at the most nine reversals. This criterion was preliminarily defined on the basis of pilot measurements. The following experiment was set up in order to be able to investigate the adaptation process more systematically.

At a constant frequency of 1 kHz the threshold was measured with FINESS for 40 reversals. The resulting thresholds are shown in Figure 9. For better comparability the thresholds are expressed relative to their mean value over the last ten reversals ('asymptotic' threshold). All thresholds clearly drop at the beginning of the measurement. The size of this drop is of minor relevance since the initial value of the shifted threshold strongly depends on the subject's absolute threshold at 1 kHz relative to the starting level of 15 dB HL. However, the number of reversals a subject needs for their threshold to settle around zero strongly varies. Some subjects reach their 'asymptotic' threshold within 4 reversals, others need more than 15 reversals. This shows that a maximum of nine reversals in the warm-up phase may not be enough to exclude all adaptation effects from the actual measurement. Also the 1.3-dB criterion between consecutive threshold values would lead to a premature break of the warm-up phase in several cases, three of which are pointed out by arrows in Figure 9. In order to improve the duration of the warm-up phase the mean threshold of all subjects was analysed (thick line in Fig. 9). It reaches a value of 0.5 dB after



**Figure 7:** Comparison of thresholds measured with FINESS (lines) and a 3-AFC procedure (circles). The original FINESS data (thin lines) were shifted vertically (thick lines) in order to minimize the sum of their squared distance to the 3-AFC data which show the mean results from two repetitions. The error bars on the 3-AFC data and the error shades on the FINESS data indicate the minimum and maximum threshold values obtained for the corresponding frequencies (cf. Fig. 2d). The bottom left corner of each panel holds the value of the correlation coefficient  $r$  between FINESS and 3-AFC data.



**Figure 9:** Normalised thresholds of 20 subjects at 1 kHz (thin lines) as a function of the number of reversals, and the mean threshold (thick line). The arrows point to examples where a criterion for ending the warm-up phase that is based on the slope of the threshold fails.

16 reversals. This is also the maximum number of reversals the subjects needed to come within  $\pm 2$  dB of their ‘asymptotic’ threshold which seems close enough to start the actual measurement. Therefore the warm-up phase in future versions of FINESS will contain a constant 16 reversals so that in the majority of cases the measurement should not be affected by the phase in which the subject gets accustomed to the method. On average the 20 subjects took 40 s ( $\pm 10$  s) for the first 16 reversals, which is about twice as much as the mean duration of the current warm-up phase (19 s) but still acceptable if by this means a bias in the threshold estimates can be largely avoided.

## 7 Discussion

The validity of the screening method FINESS and also of the FINESS-detector has been tested by a number of experiments. Most of these experiments were based on a comparison between the thresholds measured with FINESS and thresholds measured either with the same method (test/retest experiment I) or with the reference method as in experiments II & III. In these cases care was taken to conduct the measurements directly after one another in order to minimise differences in the measured thresholds due to changes in the true thresholds. The general pattern of threshold fine structure appears to be stable over a long time (Kemp, 1979). However, for SOAEs, blood flow (Long & Talmadge, 1997), body position (de Kleine et al., 2000) and adaptation to the quiet environment in a sound-treated booth (Rabinowitz & Widin, 1984; Smurzynski & Probst, 1998) have been reported to have small effects on the SOAE frequency and amplitude. Therefore, if one assumes the mechanisms underlying threshold fine structure and SOAEs to be related, similar effects on threshold fine structure are possible. Obviously such effects could not be distinguished from deviations in the thresholds which are due to the method, and would reduce the correlations in the test/retest ex-

periment.

The comparison with the reference procedure in experiment II demonstrated that in general the shape of the threshold curves is accurately estimated by FINESS. Several subjects however (e.g. HR-r, NG-l, NF-l, ZM-l) showed deviations of the depth of threshold minima (i.e. very sensitive regions of hearing). This indicates that fine structure may be slightly smoothed by FINESS, which could have various reasons. If, for instance, a subject adopts a ‘rhythmic’ response pattern, large changes in the threshold from one frequency to another will tend to be underestimated. Or — as was explained in the introduction to Section 4 — depending on the subject’s internal threshold criterion, not the absolute threshold but rather a low isophone may be measured in which the fine structure is slightly less pronounced (Mauermann et al., 2004). These are general drawbacks of tracking procedures. However, they seem tolerable when considering the speed of these procedures which enables the use of a high frequency resolution. This is essential for the measuring of fine structure, since otherwise narrow ripples may be easily overseen as e.g. is the case with the 3-AFC data of subject HA-l in Figure 7.

Experiment III revealed that only minor drifts across frequency seem to be present in the threshold curves obtained by FINESS. An exception was subject TA-l whose thresholds markedly improved towards the end of a run. Although this subject did not show unusual behaviour in experiment IV, an adaptation effect similar to but much slower than that observed in experiment IV could account for this drift. In general, however, the drifts were sufficiently small so that they should not affect the detection of fine structure, which typically has a cycle length of 1/10 octave. This effect only needs to be taken into consideration when the global shape of a threshold curve over a larger frequency range or estimates of absolute threshold curves are of interest. As the comparisons with the reference procedure in experiments II and III showed, FINESS is not able to estimate absolute thresholds very well. Nonetheless, if absolute thresholds are required, the threshold curves measured with FINESS may be adjusted by obtaining the threshold at one or two frequencies through a more bias-free method, e.g. 3-AFC. In the same way drifts may be corrected in threshold curves over larger frequency ranges.

Experiment IV led to a better characterisation of the effect of improving thresholds at the beginning of a FINESS run. This enabled an optimisation of the warm-up phase that precedes every run, so that in future the actual measurement should not be affected by this effect significantly. Yet the underlying reasons for the improving thresholds are not clear. They could be due to the subjects first having to establish their internal criterion. Or it may be an adaptation of the auditory system to the very quiet environment and the low level stimuli. Further research is needed to answer this question.

With the screening method (FINESS) and the algorithm for detecting fine structure (FINESS-detector) two tools have been introduced which allow further investigations of fine structure. Such investigations could point in several directions: 1. Fine structure and SOAEs are

thought to originate from the same resonances in the cochlea (Talmadge et al., 1998, s. Sec. 1). Whereas SOAEs are often hard to measure because of the large attenuation they receive on their way from the cochlea to the ear canal, fine structure as a perceptual consequence of the resonances may be able to give additional information and contribute to a better understanding of cochlea mechanics. 2. Studies on the occurrence of fine structure in larger populations could shed some light on the relation between fine structure and cochlear vulnerability phenomenologically. If such a relation could be established, fine structure could become a relevant indicator for early hearing loss that might be useful in clinical diagnostics. 3. Fine structure has some influence on other fields of auditory perception such as temporal integration (Cohen, 1982), loudness (Mauermann et al., 2004) and amplitude modulation (Zwicker, 1986; Long, 1993; Heise et al., 2006b). However, these issues are still not understood in detail, partly because of the work involved in measuring fine structure.

For all these studies an objective detector of fine structure should be helpful. However, different groups may prefer different aspects of the detector: when researching into the nature of fine structure it is probably beneficial to have a measure of fine structure that is local. Previous measures of fine structure only provided a global characterisation such as the summated peak height (Horst et al., 2003) or an average of the spectrum of threshold curves across the periodicity range that is typical for fine structure (Kapadia & Lutman, 1999). For clinical diagnostics on the other hand the indication of fine-structure regions or even a yes/no detector may be advantageous. The FINESS-detector returns both a measure of fine structure and regions containing fine structure. Moreover, it is implemented in a way that the criteria for fine structure may be easily adapted to meet individual needs. Because of this flexibility an application to other types of fine structure such as that of DPOAEs seems possible. Basically critical ripple depths ( $\Delta L_{\min}$ ) and periodicity ranges ( $\Delta f_{\min}$ ,  $\Delta f_{\max}$ ) may be specified from which the fine-structure measure is calculated. This measure was designed to combine several criteria, since a rigid criterion does not seem adequate for the characterisation of fine structure. This judgement is confirmed by the slight fluctuations in the depth of fine-structure ripples observed in the test/retest experiments and by the existence of a variety of criteria in the literature. Fine-structure regions may be derived from the fine-structure measure. In this paper a criterion of 3 dB<sub>avg</sub> was used for the detection of fine-structure regions, since this criterion seemed to yield the most appropriate characterisation of the threshold curves and also led to the most robust results in the test/retest experiment.

The duration of a FINESS measurement is around  $10.5 \pm 2.8$  min per octave, which is the mean duration (and the standard deviation) of all 79 measurements carried out during this study. Relatively speaking, measurements over a smaller frequency range take longer than measurements over a larger range because of the constant period for the warm-up phase. Measuring fine structure with non-tracking procedures such as standard audiome-

try, AFC or the ‘‘Audioscan’’ (Meyer-Bisch, 1996) requires at least double this time.<sup>3</sup> The speed was achieved by ‘fusing’ the fixed-frequency and the frequency sweep tracking method and using only one reversal per frequency. Compared to methods that use frequency sweeps this has two advantages: firstly, fluctuations in the reaction time of subjects do not result in frequency shifts that, in the worst case, could lead to a cancellation of the fine structure when averaging multiple runs. Secondly, a short span of inattention on the part of the subject only affects the measurement locally and not over a wider frequency range.

In summary, a fast method for measuring threshold fine structure (FINESS) and an automatic detector for locally identifying and quantifying threshold fine structure (FINESS-detector) were introduced and evaluated. By test/retest experiments and by comparison with thresholds obtained by a reference procedure it was shown that the method is sufficiently reliable, accurate and stable across frequencies to measure fine structure. The detector describes the fine structure in a threshold curve adequately and objectively, and yields reproducible results.

In the course of this study fine structure greater than 3 dB<sub>avg</sub> was found in 10 out of 22 ears (cf. Fig. 6). To draw any conclusions from this data base concerning the possible relation between fine structure and hearing loss, however, would be premature. This needs to be explored in future studies with larger populations.

## Acknowledgements

The authors would like to thank Birger Kollmeier for fruitful discussion. Parts of this study were presented at the ‘‘9. Jahrestagung der Deutschen Gesellschaft für Audiologie’’ in Cologne, Germany (Heise et al., 2006a) and at the ‘‘XXVIIIth International Congress of Audiology’’ in Innsbruck, Austria (Heise et al., 2006c). This work was supported by the Deutsche Forschungsgemeinschaft via grants GRK591/3 and KO942/18-1&2. The study was approved by the University of Oldenburg Research Ethics Committee, April 2004.

## References

- Cleveland W.S., 1979. Robust locally weighted regression and smoothing scatterplots. *J Am Stat Assoc*, 74, 829–836.
- Cohen M.F., 1982. Detection threshold microstructure and its effect on temporal integration data. *J Acoust Soc Am*, 71, 405–409.
- Elliott E., 1958. A ripple effect in the audiogram. *Nature*, 181, 1076.
- Furst M., Reshef I. & Attias J., 1992. Manifestations of intense noise stimulation on spontaneous otoacoustic emission and threshold microstructure: experiment and model. *J Acoust Soc Am*, 91, 1003–1014.
- Heise S.J., Verhey J.L. & Mauermann M., 2006a. Effizientes Verfahren zur Bestimmung der Feinstruktur

<sup>3</sup>Here an adaptation of the ‘‘Audioscan’’ method is referred to in which the frequencies are swept slow enough to detect fine structure.

- der Hörschwelle. In: *9. Jahrestagung der Deutschen Gesellschaft für Audiologie, Zeitschrift für Audiologie*. Suppl. 9, CDROM.
- Heise S.J., Mauermann M. & Verhey J.L., 2006b. Modulation-wahrnehmung nahe der Hörschwelle. In: *9. Jahrestagung der Deutschen Gesellschaft für Audiologie, Zeitschrift für Audiologie*. Suppl. 9, CDROM.
- Heise S.J., Verhey J.L. & Mauermann M., 2006c. Optimized Békésy tracking for the detection of hearing threshold fine structure. In: *XXVIIIth International Congress of Audiology*.
- Heise S.J., Verhey J.L. & Mauermann M., 2007. Ein effizientes Verfahren zur Bestimmung der Feinstruktur der Hörschwelle. *Z Audiol*, 46, 126–136.
- Horst J.W. & de Kleine E., 1999. Audiogram fine structure and spontaneous otoacoustic emissions in patients with Menière's disease. *Audiology*, 38, 267–270.
- Horst J.W., Wit H.P. & Albers F.W.J., 2003. Quantification of Audiogram Fine-Structure as a Function of Hearing Threshold. *Hear Res*, 176, 105–112.
- IEC 60645-1, 2001. *Audiological equipment – Part 1: Pure-tone audiometers*.
- ISO 389-8, 2004. *Reference zero for the calibration of audiometric equipment – Part 8: Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones*.
- ISO 8253-1, 1989. *Acoustics – Audiometric test methods – Part 1: Basic pure tone air and bone conduction threshold audiometry*.
- Kapadia S. & Lutman M.E., 1999. Reduced 'audiogram ripple' in normally-hearing subjects with weak otoacoustic emissions. *Audiology*, 38, 257–261.
- Kemp D.T., 1979. The evoked cochlear mechanical response and the auditory microstructure – evidence for a new element in cochlear mechanics. *Scand Audiol Suppl*, 9, 35–47.
- de Kleine E., Wit H.P., van Dijk P. & Avan P., 2000. The behavior of spontaneous otoacoustic emissions during and after postural changes. *J Acoust Soc Am*, 107, 3308–3316.
- Kollmeier B., Gilkey R.H. & Sieben U.K., 1988. Adaptive staircase techniques in psychoacoustics: A comparison of human data and a mathematical model. *J Acoust Soc Am*, 83, 1852–1862.
- Levitt H., 1971. Transformed Up-Down Methods in Psychoacoustics. *J Acoust Soc Am*, 49, 467–477.
- Long G.R., 1993. Perceptual Consequences of Otoacoustic Emissions. In: A. Schick (ed.) *Contributions to Psychological Acoustics: Results of the 6th Oldenburg Symposium on Psychological Acoustics*. Oldenburg, Germany: University of Oldenburg Press, pp. 59–80.
- Long G.R. & Talmadge C.L., 1997. Spontaneous otoacoustic emission frequency is modulated by heartbeat. *J Acoust Soc Am*, 102, 2831–2848.
- Long G.R. & Tubis ., 1988a. Modification of spontaneous and evoked otoacoustic emissions and associated psychoacoustic microstructure by aspirin consumption. *J Acoust Soc Am*, 84, 1343–1353.
- Long G.R. & Tubis ., 1988b. Investigations into the nature of the association between threshold microstructure and otoacoustic emissions. *Hear Res*, 36, 125–138.
- Marshall L. & Jesteadt W., 1986. Comparison of pure-tone audibility thresholds obtained with audiological and two-interval forced-choice procedures. *J Speech Hear Res*, 29, 82–91.
- Mauermann M., Long G.R. & Kollmeier B., 2004. Fine Structure of Hearing Threshold and Loudness Perception. *J Acoust Soc Am*, 116, 1066–1080.
- McFadden D. & Mishra R., 1993. On the relation between hearing sensitivity and otoacoustic emissions. *Hear Res*, 71, 208–213.
- McFadden D. & Pasanen E.G., 1994. Otoacoustic emissions and quinine sulfate. *J Acoust Soc Am*, 95, 3460–3474.
- Meyer-Bisch C., 1996. Audioscan: a high-definition audiometry technique based on constant-level frequency sweeps – a new method with new hearing indicators. *Audiology*, 35, 63–72.
- Norton S.J., Mott J.B. & Champlin C.A., 1989. Behavior of spontaneous otoacoustic emissions following intense ipsilateral acoustic stimulation. *Hear Res*, 38, 243–258.
- Rabinowitz W.M. & Widin G.P., 1984. Interaction of spontaneous oto-acoustic emissions and external sounds. *J Acoust Soc Am*, 76, 1713–1720.
- Reuter K. & Hammershøi D., 2006. Distortion product otoacoustic emission fine structure analysis of 50 normal-hearing humans. *J Acoust Soc Am*, 120, 270–279.
- Schloth E., 1983. Relation between Spectral Composition of Spontaneous Otoacoustic Emissions and Fine-structure of Threshold in Quiet. *Acustica*, 53, 250–256.
- Shera C.A. & Zweig G., 1993. *Biophysics of Hair Cell Sensory Systems*, Singapore: World Scientific, chap. Order from chaos: Resolving the paradox of periodicity in evoked otoacoustic emissions. pp. 54–63.
- Silver N.C. & Dunlap W.P., 1987. Averaging Correlation Coefficients: Should Fisher's z Transformation Be Used? *J Appl Psychol*, 72, 146–148.
- Smurzynski J. & Probst R., 1998. The influence of disappearing and reappearing spontaneous otoacoustic emissions on one subject's threshold microstructure. *Hear Res*, 115, 197–205.
- Talmadge C.L., Tubis A., Long G.R. & Piskorski P., 1998. Modeling otoacoustic emission and hearing threshold fine structures. *J Acoust Soc Am*, 104, 1517–1543.
- Zweig G. & Shera C.A., 1995. The origin of periodicity in the spectrum of evoked otoacoustic emissions. *J Acoust Soc Am*, 98, 2018–2047.
- Zwicker E., 1986. Spontaneous Oto-Acoustic Emissions, Threshold in Quiet, and Just Noticeable Amplitude Modulation at Low Levels. In: B.C.J. Moore & R.D. Patterson (eds.) *Auditory Frequency Selectivity*. New York: Plenum, pp. 49–59.
- Zwicker E. & Schloth E., 1984. Interrelation of different otoacoustic emissions. *J Acoust Soc Am*, 75, 1148–1154.