DOI: 10.61782/fa.2023.1135

forum acusticum 2023

ECHOES FROM THE ARCHIVES OF THE MUNICH SCHOOL OF PSYCHOACOUSTICS

Bernhard U. Seeber*
Audio Information Processing, Technical University of Munich, Germany

ABSTRACT

Eberhard Zwicker founded the Institute of Electroacoustics at the TH Munich, now Technical University of Munich, in 1967 and was joined by the psychoacousticians Ernst Terhardt and Hugo Fastl and, for speech recognition, by Guenther Ruske. The subsequent two+ decades were a highly fruitful period for psychoacoustics in Munich, leading to what can now be viewed as the "Munich School of Psychoacoustics": Research on peripheral auditory processes led to key contributions to spectral and temporal auditory masking and to measures of critical bandwidth (BARK bands), while work on non-linear cochlear processes, including otoacoustic emissions, resulted in a non-linear electric model of the cochlea. Terhardt's works on pitch, consonance and harmony were pioneering. Research on the percepts of sharpness, roughness and fluctuation strength laid the foundation for psychoacoustic sound quality. DIN 45631 and ISO 532b standardize Zwicker's loudness model and bridge to a wide array of applications of psychoacoustics. While I am too young to have met Eberhard Zwicker in person, I have "inherited" his personal and the institute document archives and will report on selected research outcomes.

Keywords: *Munich School of Psychoacoustics; Eberhard Zwicker; Hugo Fastl; Ernst Terhardt; Masking; Pitch; Loudness.*

****Corresponding author***: *seeber@tum.de.*

1. INTRODUCTION

The Munich Three, Eberhard Zwicker, Ernst Terhardt and Hugo Fastl, were leading researchers in basic psychoacoustics who likewise contributed with their innovations to applied acoustics and industry. Rigorous technical developments formed the basis for research and led to countless devices for measuring various acoustic and perceptual quantities: research was conducted with stringently calibrated headphones with a free-field equalizer (Figure 1 and [1]), dedicated electronic devices were developed for measuring tuning curves and masking patterns (Figure 2), and a loudness meter was able to compute and display loudness of instationary sounds. This short paper cannot give an overview of the extensive research that took place in the decades at Zwicker's Institute of Electroacoustics at the Technical University of Munich (TUM) – the research is excellently summarized in the book "Psychoacoustics – Facts and Models" [2-5]. What I am aiming at is to give pointers to a few aspects of their research that appear important to me and that impressed me as an electrical engineer who is nowadays mainly using commercial hardware and working with computers for stimulus generation and modeling.

2. AUDITORY PERIPHERAL PROCESSING AND MASKING

Understanding and modeling peripheral auditory processing was a key focus of Zwicker before coming to and then at TUM. He measured various conditions of masking: tones masking narrow-band-noise [6], tones masking tones [7], temporal masking [8], psychoacoustic tuning curves [9], masking in hearing impaired listeners [10] and later effects of binaural masking [11, 12]. Fastl also had an early research focus on temporal masking, the topic of his habilitation [13, 14].

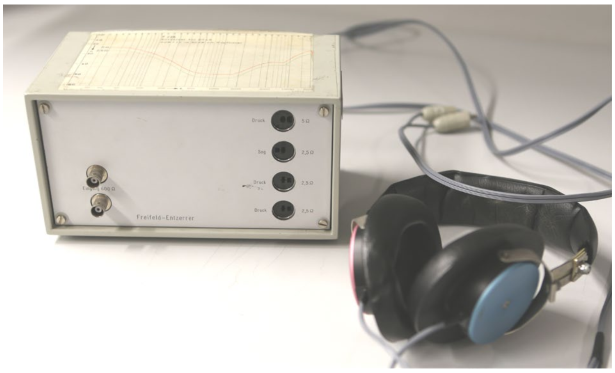

Figure 1. Custom free-field equalizer for Beyerdynamic DT48 headphones [1].

Figure 2. Custom device for measuring psychoacoustic tuning curves.

Of particular interest were the non-linear effects of masking and how they relate to loudness perception. The level-dependent non-linear growth of masking and the asymmetry of masking below and above the masker were didactically worked out in Zwicker and Jaroszewski (1982) [7] and are replotted in Figure 3. The masking difference of 20 dB between the masking pattern and the mirrored masking pattern, highlighted with "A" in Figure 1, depicts the asymmetry toward increased masking at frequencies above the masker frequency compared to below – while "B" and "C" indicate the substantial non-linear increase ("upward spread of masking") above the masker frequency with increasing masker level.

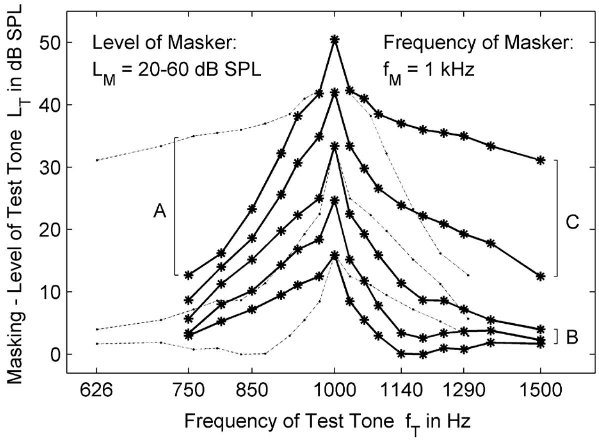

Figure 3. Tone-on-tone masking patterns are depicted with their mirrored version (on the Bark scale) to highlight the non-linearity of masking as measured by [7], replotted [c.f., 15].

Most masking effects can be understood and modeled assuming energy detection at the output of the auditory filter. This leads to the need of understanding filter bandwidth and shape. Zwicker derived the critical (filter) bandwidth measured in BARK from filter estimates obtained with several different approaches [4], amongst them the band-widening of noise when masking the probe tone, notched-noise masking and modulation-frequency dependent differences in the sensitivity to frequency and amplitude modulation. He defined the critical bandwidth in the unit BARK and an integrated critical band scale [16] which numbers the critical band filters from 1 to 24. For numerical modeling work, analytical, differentiable expressions were found for the BARK bandwidth and the critical band rate scale [17]. The critical bandwidth according to Zwicker is plotted in relation to other estimates of auditory filter bandwidth in Figure 4, as summarized in [15]. A simple estimate of the BARK bandwidth is: 100 Hz bandwidth for center frequencies up to 500 Hz and 20% of the center frequency for center frequencies above 500 Hz.

The extensive work about the auditory periphery let to a non-linear analog model of the cochlea, which was built in hardware (Figure 5). The model transmits energy between cochlear filters using a resistor network. In each cochlear filter, a phase-shifted feedback loop with non-linearity and with adjustable feedback attenuation simulates the amplification by outer hair cells (c.f. Fig. 3.9 in [4]). The

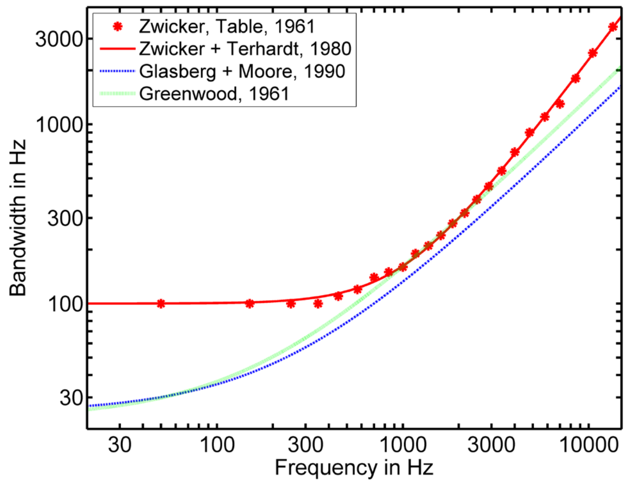

Figure 4. Measures of critical bandwidth: Zwicker's tabular data (*, [16]) and formula [17] for BARK-bandwidth, the ERB-bandwidth by Glasberg and Moore [18], and the physiological bandwidth after Greenwood [19]. Replotted after [15].

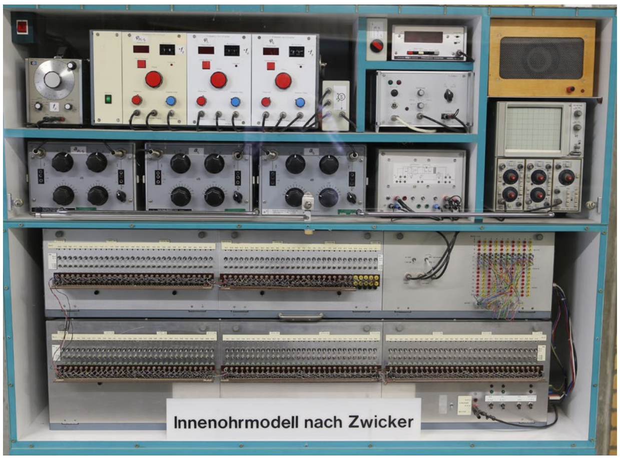

Figure 5. Hardware implementation of the non-linear inner ear model of Zwicker [3]. The filterbank (resolution 0.1 BARK) with non-linear feedback and connection across neighboring filters is visible at the bottom. On the top, two generators for tones with frequencies f1 and f2 are visible, and a third one with a phase shifter to generate a tone at 2f1-f2 to cancel the distortion product ("otoacoustic emission") created by the non-linearity of the inner ear model. The output of each filter can be inspected on the oscilloscope and listened to.

distortion products were measured by phase-true cancellation [3]. The model was used for research on distortion products created by the inner ear and otoacoustic emissions.

3. LOUDNESS

Given Zwicker's extensive knowledge about masking and the auditory periphery, the loudness of sounds was of paramount interest. Several studies investigated how loudness develops above masked threshold and integrates across critical bands [20], leading to the loudness model by Zwicker, which is based on collaborative work with Stevens and Scharf [20-22]. The model was developed into the ISO 532B standard [23] and later extended for dynamic sounds [24] and for hearing impairment [25]. While the early standard used a paper chart to compute loudness, digital hardware implementations were soon available for use in technical acoustics settings.

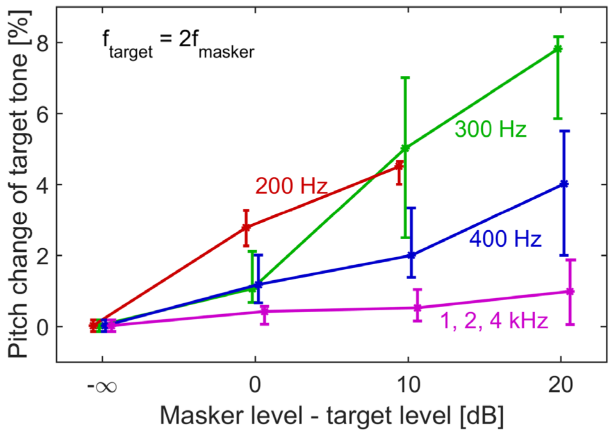

Figure 6. Pitch shift of a target tone twice in frequency of the masker tone, as a function of its relative level to the masker. Pitch shifts up by up to 8% for partially masked tones. Replotted after [26].

4. PITCH, CONSONANCE AND HARMONY

Ernst Terhardt's focus was on pitch perception and music. In psychoacoustic experiments, he measured pitch shifts to find the effects of level and partial masking by other spectral components. For example, the pitch of a spectral component can shift up (or down) by several percent if it is partially masked (Figure 6, [26]). These masking effects also extend to harmonic complexes and agree with the pitch being derived from a spectral, place pitch mechanism.

Based on his findings, he developed a model of virtual pitch based on subharmonic coincidence matching, which predicts various virtual pitch phenomena, including pitch ambiguities and changes in virtual pitch due to masking and component level [26, 27]. His insights are summarized in the book "Akustische Kommunikation" [28].

5. TEMPORAL PROCESSING AND SOUND QUALITY

Hugo Fastl likewise worked on basic psychoacoustics, with a focus on temporal processing [13, 14, 25], but he is also known for translating psychoacoustic knowledge to technical acoustics. He advocated the use of psychoacoustic quantities in the sound design process [29] and worked on composite sound quality measures, e.g. psychoacoustic annoyance/pleasantness derived from factors of loudness, sharpness, roughness and fluctuation strength. Several studies at the institute investigated roughness, which was found to follow a bandpass-characteristic irrespective of level and masking: roughness increases from low modulation rates to peak at around 70 Hz modulation rate, to decrease thereafter at higher modulation rates [30], see Figure 7.

His interest in temporal processing is also visible in his suggestion for a fluctuating noise for speech audiometry, now known as "Fastl-noise" [31]. Speech intelligibility strongly depends on our ability to listen into temporal gaps in a temporally modulated noise masker, like speech. He suggested a noise that not only has the long-term spectrum of speech but also shares the modulation spectrum with speech. Since it is noise, it will not contain any meaning and thus not cause "informational" masking.

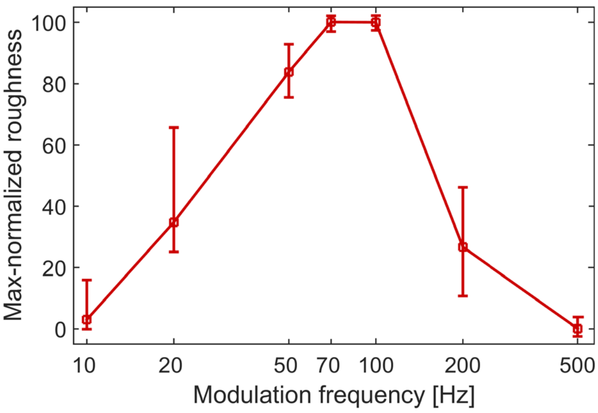

Figure 7. The roughness of amplitude modulated tones. Replotted after [30].

Figure 8. Zwicker Institute publication archive containing pre-prints and background material on almost 800 publications of the institute.

6. ANNOTES

As hinted at in the introduction, the research approach at Zwicker's institute was of thorough engineering: custom devices were built when necessary and deep understanding of electrical engineering and acoustics was assumed. The organization was thorough as well and is still leading in today's times of open science data management: the publication archive contains background material to almost 800 publications, including pre-prints, data, notes and figure drawings (Figure 8).

7. REFERENCES

[1] E. Zwicker, and D. Maiwald, "Über das Freifeldübertragungsmaß des Kopfhörers DT 48," *Acustica,* vol. 13, pp. 181-182, 1963.

[2] E. Zwicker, and R. Feldtkeller, *Das Ohr als Nachrichtenempfänger*, Stuttgart: S. Hirzel Verlag, 1967.

[3] E. Zwicker, and H. Fastl, *Psychoacoustics, Facts and Models*: Springer, 1990.

[4] E. Zwicker, and H. Fastl, *Psychoacoustics, Facts and Models*, 2. ed., Berlin, Heidelberg, New York: Springer, 1999.

[5] H. Fastl, and E. Zwicker, *Psychoacoustics - Facts and Models*, 3. ed., Berlin, Heidelberg: Springer, 2007.

[6] E. Zwicker, "Die Verdeckung von Schmalbandgeräuschen durch Sinustöne," *Acustica,* vol. 4, pp. 415-420, 1954.

[7] E. Zwicker, and A. Jaroszewski, “Inverse frequency dependence of simultaneous tone-on-tone masking patterns at low levels,” *J. Acoust. Soc. Am.,* vol. 71, no. 6, pp. 1508-1512, 1982.

[8] E. Zwicker, “Temporal effects in simultaneous masking by white-noise bursts,” *J. Acoust. Soc. Am.,* vol. 37, pp. 653-663, 1965.

[9] E. Zwicker, "On a psychoacoustical equivalent of tuning curves," *Facts and Models in Hearing*, E. Zwicker and E. Terhardt, eds., pp. 132-141: Springer, 1974.

[10] E. Zwicker, and K. Schorn, “Psychoacoustical Tuning Curves in Audiology,” *Audiology,* vol. 17, pp. 120-140, 1978.

[11] E. Zwicker, and G. B. Henning, “Binaural masking-level differences with tones masked by noises of various bandwidth and levels,” *Hearing Research,* vol. 14, pp. 179-183, 1984.

[12] E. Zwicker, and G. B. Henning, “The four factors leading to binaural masking-level differences,” *Hearing Research,* vol. 19, pp. 29-47, 1985.

[13] H. Fastl, “Temporal masking effects: II. Critical band noise masker,” *Acustica,* vol. 36, pp. 317-330, 1976.

[14] H. Fastl, “Temporal masking effects: III. Pure tone masker,” *Acustica,* vol. 43, pp. 282-294, 1979.

[15] B. Seeber, "Masking and Critical Bands," *Handbook of Signal Processing in Acoustics*, D. Havelock, S. Kuwano and M. Vorländer, eds., pp. 229-240, Berlin, New York: Springer, 2008.

[16] E. Zwicker, “Subdivision of the audible frequency range into critical bands (Frequenzgruppen),” *J. Acoust. Soc. Am.,* vol. 33, no. 3, pp. 248, 1961.

[17] E. Zwicker, and E. Terhardt, “Analytical expressions for critical band rate and critical bandwidth as a function of frequency,” *J. Acoust. Soc. Am.,* vol. 68, no. 6, pp. 1523-1525, 1980.

[18] B. R. Glasberg, and B. C. J. Moore, “Derivation of auditory filter shapes from notched-noise data,” *Hearing Research,* vol. 47, no. 5, pp. 103-138, 1990.

[19] D. D. Greenwood, “Critical Bandwidth and the Frequency Coordinates of the Basilar Membrane,” *J. Acoust. Soc. Am.,* vol. 33, no. 10, pp. 1344-1356, 1961.

[20] E. Zwicker, G. Flottorp, and S. S. Stevens, “Critical band width in loudness summation,” *J. Acoust. Soc. Am.,* vol. 29, no. 5, pp. 548-557, 1957.

[21] E. Zwicker, and B. Scharf, “A Model of Loudness Summation,” *Psych. Rev.,* vol. 72, pp. 3-26, 1965.

[22] E. Zwicker, “Über psychologische und methodische Grundlagen der Lautheit,” *Acustica,* vol. 8, pp. 237-258, 1958.

[23] ISO, "ISO Recommendation R 532: Method for calculating loudness level," Int. Standardisation Organisation, 1975.

[24] DIN, "DIN 45631/A1: Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum - Verfahren nach E. Zwicker - Änderung 1: Berechnung der Lautheit zeitvarianter Geräusche; mit CD-ROM," 2010.

[25] J. Chalupper, and H. Fastl, “Dynamic loudness model (DLM) for normal and hearing-impaired listeners,” *Acta Acustica - Acustica,* vol. 88, no. 3, pp. 378-386, 2002.

[26] E. Terhardt, and H. Fastl, “Zum Einfluss von Störtönen und Störgeräuschen auf die Tonhöhe von Sinustönen,” *Acustica,* vol. 25, pp. 53-61, 1971.

[27] E. Terhardt, “Calculating virtual pitch,” *Hearing Research,* vol. 1, no. 2, pp. 155-182, 1979.

[28] E. Terhardt, *Akustische Kommunikation: Grundlagen mit Hörbeispielen*, Berlin, Heidelberg: Springer, 1998.

[29] H. Fastl, "Noise Evaluation Based on Hearing Sensations," *Proc. of the 7th Western Pacific Regional Acoustics Conference, WESTPRAC VII*, The Acoustical Society of Japan, 2000, pp. 33-41.

[30] H. Fastl, “The hearing sensation roughness and neuronal responses to AM-tones,” *Hear Res,* vol. 46, no. 3, pp. 293-295, 1990.

[31] H. Fastl, “Ein Störgeräusch für die Sprachaudiometrie (A background noise for speech audiometry),” *Audiologische Akustik,* vol. 26, no. 1, pp. 2-13, 1987.