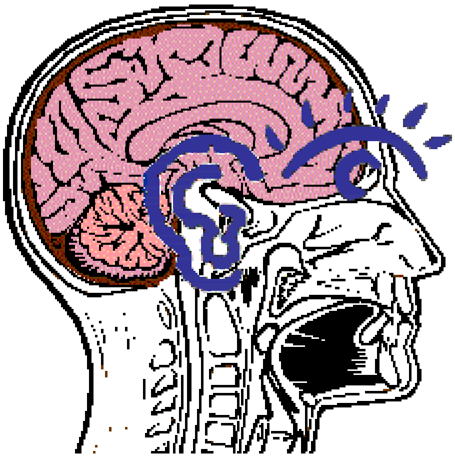
CARL VON OSSIETZKY universität OLDENBURG

European Graduate School for



Euro-GK

Neurosensory Science, Systems and Application

Summerschool and Symposium

"*Psychophysics, physiology and models of the central auditory system*"

12.08.2001 - 16.08.2001
in Bad Zwischenahn

Bildungsstätte der Angestelltenkammer Bremen, Am Rosenteich 26, in D-26160 Bad Zwischenahn

Organisation:
Prof. Dr. Dr. B. Kollmeier
AG Medizinische Physik, FB 8 Physik
z.H. Ingrid Wusowski
Universität Oldenburg
D - 26111 Oldenburg
Tel.:0441-798-5470
Fax.:0441-798-3902
e-Mail: ingrid.wusowski@uni-oldenburg.de

Program

Sunday, Aug. 12, 20.00 Get-together-Dinner

Monday, August 13th " Binaural Processing "

09.00 - 09.15	Birger Kollmeier	Welcome and Introduction
09.15 - 10.05	Dorte Hamershoij	Binaural technique
10.05 - 10.40		Coffee-break
10.40 - 11.30	Alan Palmer	Physiological basis of the BMLD
11.30 - 12.20	Yoichi Suzuki	Auditory search asymmetry under a spatial listening condition
12.20 - 14.00		Lunch-break
14.00 - 17.15		Poster session
17.15 - 18.05	Armin Kohlrausch	A unified view on spectral integration in binaural hearing.
18.15 - 19.15		Dinner
20.00 - 21.30		Evening Discussion: Hearing Research - Quo vadis ?

Tuesday, August 14th " Spectrotemporal Processing and Masking "

09.00 - 09.50	Neal Viemeister	Intensity discrimination, loudness, and internal noise
09.50 - 10.40	Laurel Carney	Physiologically-based models for masking
10.40 - 11.10		Coffee-break
11.10 - 12.00	Christopher Plack	Linearity and nonlinearity in masking
12.00 - 12.15		Discussion
12.15 - 14.00		Lunch-break
14.00 - 17.15		Workshop (Group work)
17.30 - 18.15	Ray Meddis	Signal processing in the cochlear nucleus
19.00 - ??		Boat trip with Ammerländer Eel Dinner

Wednesday, August 15th " Auditory object perception **"**

09.00 - 09.50	Brian Moore	Factors influencing sequential stream segregation
09.50 - 10.40	Holger Schulze	Representation of stimulus periodicity and its learning induced plasticity in the auditory cortex
10.40 - 11.10		Coffee-break
11.10 - 12.00	Leo v. Hemmen	Zwicker tone and noise reduction: A neurophysical model based on double negatives
12.00 - 12.15		Discussion
12.15 - 14.00		Lunch-break
14.00 - 14.50	Shihab Shamma	The representation of Timbre and Pitch in the Auditory Cortex
14.50 - 15.20		Coffe break
15.20 - 16.10	Georg Klump	A comparative view on CMR
16.10 - 17.00	Bob Carlyon	The Continuity Illusion: Vowel identification, Frequency Modulation, and Hearing Backwards in Time
18.00 - 19.30		Dinner
20.30 - ???		An evening in the Casino (please sign on and bring a tie and a passport!)

Thursday, August 16th " Speech perception and processing **"**

09.00 - 09.50	Tammo Houtgast	Psychoacoustics and speech reception by the hearing impaired
09.50 - 10.40	Arne Leijon	Estimation of auditory information transmission capacity using a hidden Markov model of speech stimuli
10.40 - 11.10		Coffee-break
11.10 - 12.00	Hynek Hermansky	(to be announced)
12.00 - 12.30		Final Discussion
12.30 - 13.30		Lunch-break
14.00		End of meeting

Acknowledgement

The organizers want to thank the Deutsche Forschungsgemeinschaft for supporting the summerschool and symposium within the framework of the „Europäisches Graduiertenkolleg Neurosensorik“ and the „Sonderforschungsbereich Neurokognition“.

Further financial support is provided by the companies Widex (Vaerloese, Denmark), Phonak (Stäfa, Switzerland), Siemens Audiologische Technik (Erlangen, Germany), Bernafon (Berne, Switzerland) and GNReSound (Taastrup, Denmark) which is gratefully acknowledged.

Special thanks to all members of the „Arbeitsgruppe Medizinische Physik“ and of the „Europäisches Graduiertenkolleg Neurosensorik“ who contributed to organizing the meeting!

Monday, August 13th, 09:15 – 10:05

Binaural Technique

Dorte Hammershøi

Department of Acoustics, Aalborg University, Denmark

The term binaural technique is used a headline for methods for sound recording, synthesis and reproduction, where the sound pressure signals recorded (or synthesized) are the eardrum signals of the listener, and where successful reproduction is achieved, if these are truely reproduced in their original form. This lecture will review the results of a series of investigations that are the current basis for the utilization of binaural technique at Aalborg University and possible elsewhere, including investigations on sound transmission in the ear canal, measurements of HRTFs (head-related transfer functions), calibration of headphones and localization experiments in real life and with binaural recordings from real heads as well as artificial heads.

Monday, August 13th, 10:40 – 11:30

The Physiological Basis of the BMLD

A.R. Palmer and T.M. Shackleton

MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, UK

The binaural masking level difference (BMLD) was discovered in 1946 by Licklider and more fully described in the same year by Hirsh. In essence it is the greater detectability of signals, in the presence of a masker, that accrues as a result of binaural processing. While it was discovered using speech signals masked by a broadband noise, it has been extensively investigated in a much simpler form: a single low-frequency tone masked by noise. When the noise is the same at both ears, a tone presented 180° out-of-phase at the two ears is 10-15 dB more detectable than when the tone is identical at the ears. The attraction of this psychophysically measured effect is firstly, that it is very large and robust and secondly, that it can be thought of as a surrogate for our ability to better detect signals when they arrive from a different spatial location from masking sounds. Sophisticated computer models by Colburn and later by Stern and others, based upon interaural cross correlation which is assumed to occur in brainstem coincidence detectors, have provided a good description of the mechanisms underlying the BMLD.

Surprisingly, given its size and the extensive psychophysical literature that describes the BMLD it has received relatively modest attention in physiological experiments. In the earliest experiments, BMLD-generating stimuli were presented and recordings were made in the medial superior olive: these, and a few later experiments in inferior colliculus, were inconclusive. More recently, in our laboratory we have conducted a series of investigations which have verified the major tenets of the cross correlation models of BMLD.

In our first experiments, we measured responses to tones at the best frequency of the neurone and with an interaural delay at which the discharge rate was maximal, as a function of the interaural delay of the masking noise. With the signals and masker optimised for each unit's delay sensitivity, an adaptive threshold tracking procedure (mimicking the paradigms used psychophysically) detecting discharge rate increases was used to determine masked threshold for the different stimulus configurations. These results were not easily interpretable in terms of the more usual BMLD paradigms, but certainly indicated that single neurones were sensitive to the parameters of BMLD stimuli. Later, still using optimised signals and threshold tracking, but with identical or inverted noise maskers, we measured masking level differences generally in a direction, and of a magnitude, consistent with appropriate psychophysical observations in human subjects.

Next we used the more conventional BMLD stimuli with a tone fixed at 500 Hz, but with the noise level set with respect to each neural threshold. In this study, we wanted to see how the responses to BMLD stimuli related to other neural response characteristics such as the delay sensitivity. We used a method that allowed us to make a d′ computation to obtain threshold which allowed us to assess both increases and decreases in response to the BMLD signals. This proved decisive, since the cross-correlation models suggest that it is an asymmetry between increased discharge to in-phase signals and decreased discharge to out-of-phase signals that underlies the greater detectability of the latter. This was exactly what we found. The most sensitive indicator for the presence of the out-of-phase tone was indeed a decrease discharge. The in-phase tone generally caused a discharge increase, but at higher sound levels. Further experiments confirmed that the spike rate changes were a result of the effect of the tones on the interaural correlation of the noise, which is a fundamental property of the cross correlation models of the BMLD.

Finally, we adopted a population approach in which we obtained masked thresholds from as many neurones as possible in response to the simplest BMLD configurations to estimate the "population response" to a 500 Hz tone in a fixed level masker noise. We were able to show that across the population detection of the tones is by neurones with best frequencies near the tone frequency and that it depended upon discharge rate decreases in the out-of-phase condition and on increases for the in-phase condition.

Monday, August 13th, 11:30 – 12:20

Large Auditory search asymmetry under a spatial listening condition

Yoiti Suzuki, Noriaki Asemi
(Res.Inst.Electr.Commu./Grad.Sch.Info.Sci., Tohoku Univ., Japan)
Yoichi Sugita (Neuroscience Research Institute, AIST, Japan)

Auditory search asymmetry was examined under a spatial listeningcondition. The response time to detect a target sound amongdistracting sound(s) was measured. In the first experiments, results show that the response time to detect a narrow-band noise, AM tone and FM tone among distracting pure-tones was hardly affected by the number of the distractors, while the time required to detect a pure-tone among distractive narrow-band noises, AM tones and FM tones increased with the number of distractors. In the second experiment, results showed that the response time to detect a normal speech among distractive time-inverted speeches was hardly affected by the number of distractors, while the time required to detect a time-inverted speech was increased with the number of distractors. The former results suggest that our auditory system utilizes temporal changes in amplitude and frequency of sound as a basic feature for the detection of a sound in a sound-environment, the latter results suggest that the phrazes meaning to our communication tend to be recognized more quickly than

meaningless phrazes. Moreover, some fMRI images during the search task will be also introduced and discussed.

Monday, August 13th, 17:15 – 18:05

A unified view on spectral integration in binaural hearing

Armin Kohlrausch, Jeroen Breebaart, Steven van de Par

The concept of the critical band is fundamental for the evaluation of masked thresholds in monaural experiments. Both the masking of sinusoidal signals by bandpass and by bandstop noise maskers can be explained by the assumption that the masked threshold of a narrow-band stimulus is only dependent on the masker power within a certain frequency range around the test frequency. For binaural listening conditions, the estimates of the critical bandwidth are much more diverse. For some experimental procedures, like for bandstop noise maskers, the estimates are the same as for monaural conditions. For bandpass noise maskers, however, the estimates of the binaural critical bandwidth are generally larger by a factor 2 to 3 than the monaural estimates.

In this talk, we will discuss, how one can account for these experimental observations by making three assumptions aboutbinaural detection thresholds.
-- First, the detection threshold is limited by internal limits of accuracy rather than by external stimulus fluctuations.
-- Second, the binaural auditory system is able to integrate information about the presence of a signal across several auditory filters.
-- Third, the internal errors are independent in each auditory filter.

A binaural signal detection model including these assumptions correctly predicts the apparantly different binaural critical bandwidths in the various experimental conditions. The reasoning behind this explanation of the binaural phenomenon is also supported by another (monaural) experiment. If a monaural band-limiting experiment is performed with frozen-noise samples instead of running noise (i.e., there is no stimulus uncertainty), the thresholds also reveal a critical bandwidth which is larger than the estimate obtained with running noise. These results strongly suggests that internal errors of processing and external stimulus uncertainty are two important factors which lead to different apparent critical bandwidths.

Tuesday, August 14th, 09:00 – 09:50

Intensity discrimination, loudness, and internal noise

Neal Viemeister

Dept. of Psychology, Univ. of Minnesota, Minneapolis, Minnesota, USA

The relationship between differential sensitivity to intensity and the growth of subjective magnitude is a long-standing and fundamental issue in psychology, dating back to Fechner’s classic work in the mid -1800’s. In audition the issue becomes that of relating intensity discrimination (and AM detection) to loudness growth. This issue has received renewed attention because of the recent interest in the role of cochlear nonlinearities, specifically compression, in basic psychophysical situations, including loudness growth. This paper reviews recent attempts to relate intensity discrimination and loudness, and presents analyses that suggest that if there is a relationship, it is very unclear. The major uncertainty is the role of “internal noise” the intrinsic variability in the response to stimulation that limits, and perhaps determines, performance in intensity discrimination but which does not appreciably affect loudness growth. It is clear that by postulating certain properties of internal noise one can quantitatively relate the basic data on intensity discrimination, the “Weber function”, to the data on loudness growth and loudness matching. However, the derived or assumed properties of internal noise appear to contradict the physiological data and, in general, such efforts appear to provide little theoretical insight. The conclusion is that intensity discrimination may be related to loudness growth but only indirectly. Intensity discrimination is likely based on changes in local excitation, whereas loudness reflects a more global response. If so, then loudness may be of secondary importance in perception.

Tuesday, August 14th, 09:50 – 10:40

A Physiologically-Based Model for Masking

Laurel Carney

Boston University, 44 Cummington Street, Boston, MA 02215, USA

We present a model for tone-in-noise detection based on a physiologically realistic mechanism for processing the information in neural discharge times, which provides an alternative to the classical critical-band energy model. The model exploits the frequency-dependent phase properties of the tuned filters in the auditory periphery and uses across-auditory-nerve-fiber coincidence detection to extract temporal cues. The responses of these coincidence detectors are reduced when a low-frequency tone is added to a noise due to the phase opponency inherent in particular pairs of fibers tuned to different frequencies. This model predicts the detectability of low-frequency tones in roving-level noise, a psychophysical task for which the classical energy model fails.

Tuesday, August 14th, 11:10 – 12:00

Linearity and non-linearity in masking

Christopher Plack

Department of Psychology, University of Essex,
Wivenhoe Park, Colchester, CO4 3SQ, UK.

In the past it was often convenient to assume that the characteristics of spectral and temporal masking were determined by linear mechanisms: the auditory filter and the temporal window respectively. It is clear, however, that this assumption can lead to severe inaccuracies when

modelling psychophysical performance, especially in situations in which the signal and masker are of different levels and/or frequencies. For the purpose of producing a realistic model, it is important to determine the source or sources of the underlying non-linearity. In this presentation, evidence will be provided that the main non-linearity with regard to masking is the response of the basilar membrane, and that after cochlear filtering the auditory system may be regarded as linear in the intensity domain. Based on this useful property, it will be shown how non-simultaneous masking experiments may provide information about the response characteristics of the basilar membrane that is not derived easily from physiological measurements (for example, the response at low to medium characteristic frequencies).

Tuesday, August 14th, 17:30 – 18:15

Signal processing in the cochlear nucleus

Ray Meddis

Centre for the Neural Basis of Hearing, Department of Psychology,
University of Essex, Colchester, UK,

The cochlear nucleus (CN) is the first neural relay station between the auditory periphery and the cortex. As such, it might be thought unlikely to be responsible for the complex signal processing that appears to be involved in many psychophysical measurements. Nevertheless, evidence is accumulating that the opposite may be the case. Physiological studies have suggested that psychophysical judgements such as, CMR, forward masking, pitch discrimination, precedence and segregation of simultaneous sounds, have parallels in the response of single units in the CN. Accumulating evidence concerning complex anatomical structure of the CMR provide further evidence that the necessary infrastructure is present at a very low level in the brainstem to support complex signal processing. The talk will review some of this evidence and present working models that illustrate some of the principles.

Wednesday, August 15th, 09:00 – 09.50

Factors Influencing Sequential Stream Segregation

Brian C.J. Moore
Department of Experimental Psychology, University of Cambridge,
Downing Street, Cambridge CB2 3EB, U.K.
Hedwig Gockel
CNBH, Department of Physiology, University of Cambridge, Downing Street, Cambridge
CB2 3EG, U.K.

This paper reviews the factors that influence streaming, i.e. whether a sequence of sounds is heard as emanating from a single source (called fusion or coherence) or from more than one source (called fission or stream segregation). A traditional view is that streaming depends on “peripheral channelling”, the degree to which successive sounds excite different “channels” either across frequency or across ears. However, a body of evidence suggests that other factors can play a strong role; stream segregation can occur when successive sounds are presented to the same ear and have essentially identical excitation patterns. The other factors include: differences in envelope; differences in fundamental frequency; differences in phase spectrum (which can affect the waveform and hence the pitch and subjective quality of sounds); and differences in lateralisation produced by interaural time or intensity differences. Furthermore, stream segregation produced by these other factors can be obligatory, i.e. it occurs even in tasks where better performance would be achieved if successive sounds could be fused. To account for these findings, we propose the following hypothesis: the extent to which sequential stream segregation occurs is directly related to the degree of perceptual difference between successive sounds. *Any* sufficiently salient perceptual difference may lead to stream segregation, regardless of whether or not it involves peripheral channelling.

Wednesday, August 15th, 09:50 – 10.40

Representation of stimulus periodicity and its learning induced plasticity in the auditory cortex

Holger Schulze

Leibniz Institute for Neurobiology, Brenneckestr. 6, 39118 Magdeburg, Germany

The perceptual quality associated with periodic 100% sinusoidal amplitude modulated tones (AM) varies as a function of modulation frequency (f_m): AM of low f_m (up to about 100 Hz) evoke percepts of rhythm and roughness and those of higher f_m percepts of periodicity pitch. A recent study provided evidence that these different perceptual qualities are paralleled by differences in neuronal responses [1]: AM of low f_m are represented by a temporal, non-topographical (synchrony) code in primary auditory cortex (AI) and AM of high f_m by a spatial (rate-place) code (periodicity map). Furthermore, learning performance also differs for low and high f_m in an AM discrimination task in gerbils: Learning proceeds faster and discrimination performance is slightly better for low than for high f_m [2]. Based on these latter results we follow the hypothesis that different learning mechanisms might be involved in discrimination learning of AM with different low or high f_m, that is, the synchrony or rate-place representation might be altered by different learning mechanisms. Whereas learning induced alterations of stimulus representations in sensory maps are well described (e.g. [3]) only little is known about the plasticity of temporal stimulus representations like the synchrony code.

By mapping of the AI before and after AM discrimination learning in individual animals we show that the number of neurons that code for the periodicity of the training stimuli increases with learning, and that specifically those neurons change their periodicity tuning which are also tuned to the spectral content of the training stimuli. Furthermore, by recording from implanted electrodes, responses from individual neurons or small clusters of neurons could be obtained before and after the training, revealing plastic changes of temporal and spectral response properties in AI neurons that were both stimulus related and unrelated. The data suggest a neural mechanism of differential alteration of the timing of excitatory and inhibitory afferent input to auditory cortical neurons due to AM discrimination learning.

[1] Schulze H, Langner G, 1997, J Comp Physiol A 181:651-663
[2] Schulze H, Scheich H, 1999, Neurosci Lett 261:13-16
[3] Recanzone GH et al., 1993, J Neurosci 13:87-103

Wednesday, August 15th, 11:10 – 12.00

Zwicker tone and noise reduction: A neurophysical model based on double negatives

Leo v. Hemmen

Abstract not available

Wednesday, August 15th, 14:00 – 14.50

The representation of Timbre and Pitch in the Auditory Cortex

Shihab A. Shamma

Center for auditory and acoustic research, Institute for Systems research
Electrical and Computer Engineering
University of Maryland, College Park MD 20742

To understand the representation of broadband, dynamic sounds in Primary Auditory Cortex (A1), we characterize its responses by the Spectro-Temporal Response Field (STRF). The STRF describes and predicts the linear response of neurons to sounds rich with spectro-temporal envelopes. It is calculated from responses to broadband sounds with rippled spectral envelopes that drift up and down the frequency axis at various speeds. These stimuli allow us also to compute a "ripple transfer function" which summarizes the way a cell responds to all ripples. In this talk, we shall first summarize how the transfer function relates to the STRF, how it can be used to investigate the spectral and temporal response properties of the cell, and what implications these properties have to the connectivity of the cell within the cortex, and to the thalamus. We shall also address the functional implications of these results to the processing of complex sounds such as speech and music, and the relationship between auditory and other sensory processing in the cortex.

Wednesday, August 15th, 15:20 – 16.10

A comparative view on comodulation masking release

Georg M. Klump

Institute for Zoology, Technical University Munich, Lichtenbergstr. 4,
D-85747 Garching, Germany

Background noise in the natural environment constrains acoustic communication. Vertebrate auditory systems, however, have evolved efficient mechanisms to exploit temporal amplitude fluctuations of the noise to improve signal detection. This effect, that was first described in humans, has been termed comodulation masking release (CMR, see Hall et al., J Acoust Soc Am 76: 50-56, 1984). In the past years, various aspects of CMR have been studied in bird and mammalian species both psychophysically and physiologically. The presentation will evaluate and summarize these data and compare psychophysical and physiological performance.

Furthermore, the role of differences between birds and mammals in the coding of the acoustic environment will be discussed.

Wednesday, August 15th, 16:10 – 17.00

The Continuity Illusion: Vowel Perception, Frequency Modulation, And Hearing Backwards In Time

R.P. Carlyon, J.M. Deeks, D. Norris, and S. Butterfield

MRC Cognition & Brain Sciences Unit, 15 Chaucer Rd., Cambridge CB2 2EF, England.

When a "target" sound is turned off and then resumed a short time later, it can be heard as continuous, provided that the silent interval is filled by another sound that would have masked the target if it had actually remained uninterrupted. Hence both the level and frequency content of the "inducing sound" are crucial. We performed a series of experiments investigating this "continuity illusion" and its relationship to other aspects of auditory processing. In one study, we generated four different vowels, each consisting of two formants (F1 and F2). When the two formants were presented simultaneously, identification performance was very good. In a second condition, they were alternated for one second, so that F1 and F2 were never present at the same time; the duration of each formant presentation was 100 or 200 ms. Performance in this condition was close to chance. In a third condition, the F1s and F2s still alternated, but the silent intervals in each formant region were filled by noise bursts. The same noise burst was used to fill the gaps for all the F2s used, and its level was set in a preliminary experiment to induce the illusion of continuity for all F2s presented in isolation, and to fail to do so for all F1s. Similarly, the noise used to fill all F1 gaps induced continuity for all F1s in isolation, but for no F2s. Performance in this condition was substantially better compared to the condition with no noise, and to other conditions in which noise was added only to the F1 or F2 gaps. This demonstrates that the neural mechanisms responsible for vowel perception receive input from those underlying the continuity illusion. A second study investigated the finding that, when a frequency modulated (FM) tone is interrupted, and that interruption filled by noise, listeners not only hear the tone as continuous, but also hear the modulation continue through the noise. We wondered whether the *phase* of FM would be preserved during the illusion. To test this, we asked subjects to discriminate between two stimuli, both of which consisted of two portions of a 1-kHz tone modulated at a rate of 5 Hz, and separated by a 200-ms interval filled by noise. The level and frequency content of the noise were sufficient to induce the continuity illusion. In one of the two sounds the FM phase was the same after the noise as it would have been if the tone had been uninterrupted. Subjects could not discriminate between this sound and one in which the FM phase after the noise was shifted by 180°. This shows that FM phase is not preserved in the illusion, and demonstrates a paradoxical percept: subjects hear a modulation as continuous, but do not notice what would be an obvious phase reversal in that modulation. Finally, we presented listeners with a 300-ms wideband noise, which was immediately followed (without interruption) by a 300-ms narrowband noise. When asked to adjust the duration of a second narrowband noise presented 500 ms later, they adjusted it to a duration of about 370 ms. This is consistent with the onset of the first narrowband noise being perceived as occurring before the end of the wideband noise. We will present additional data investigating this explanation. If correct, it is an example of "hearing backwards in time": a subsequent sound (narrowband noise) affects what is heard before the end of a preceding sound (wideband noise).

Thursday, August 16th, 09:00 – 09.50

Psychoacoustics and speech reception by the hearing impaired

T. Houtgast

Dept. of Otolaryngology, University Hospital VU,
P.O. Box 7057, 1007 MB Amsterdam, The Netherlands
Also: TNO Human Factors, P.O. Box 23, 3769 ZG Soesterberg, The Neherlands

Sensorineural hearing impaired listeners often suffer from a reduced speech reception in noise. Detailed studies show that this group cannot be considered as homogeneous. For some, speech reception is essentially normal when taking their raised hearing threshold into account. For others, speech reception is worse than expected on the bases of their raised hearing threshold alone, suggesting additional effects of ‘supra-threshold deficits’.
Over fifty subjects with sensorineural hearing loss were subjected to SRT-tests in the presence of various types of noise (SRT is the Speech Reception Threshold, the speech-to-noise ratio for 50% sentence intelligibility),. We applied the SII-model (ANSI S3.5-1997) to predict the SRTs for each subject, accounting for the individual hearing thresholds. A deviation of the actually measured SRT from the SII-predicted SRT was taken as a measure for the effect of supra-threshold deficits on the SRT. For about half of the subjects, measured and predicted SRTs were equal (no supra-threshold deficits). For the other half, the measured SRTs were substantially poorer than predicted, suggesting the effect of supra-threshold deficits. In identifying the nature of these deficits, additional psychoacoustic measurements were performed with these subjects. The results indicate that for the listeners with a supra-threshold deficit, this deficit is related to either poor frequency resolution or to poor temporal resolution. Both aspects in one subject were observed only occasionally.

Thursday, August 16th, 09:50 – 10.40

Estimation of sensory information transmission using a hidden Markov model of speech stimuli

Arne Leijon

KTH-Speech, Music and Hearing
Drottning Kristinas v. 31, SE-100 44 Stockholm

A method is presented which gives good approximate estimates of the rate of information (in bits/s) successfully transmitted from a speech source to the modelled neural output of the peripheral sensory system. This information rate sets definite upper limits on the listener's speech-recognition performance. The performance limits depend on the entropy and vocabulary size of the speech material. The method requires a recording with clean speech and noise on separate channels, but has no restrictions on the model of sensory processing.

The estimates of sensory information rate can be used to evaluate to what extent a listeners' performance is limited by peripheral loss of information or by suboptimal central processing. Calculations for a Swedish word-recognition test material, with an excitation-pattern auditory model, agree with speech recognition results in speech-shaped masking noise. This suggests that the scarcity of sensory information may be the primary limiting factor in this test condition. Similar calculations for low-pass- and high-pass-filtered clean speech indicate a higher sensory information rate than required for the listeners' actual performance. These results suggest that the speech recognition performance under masking and filtering may be limited by different mechanisms.

Thursday, August 16th, 11:10 – 12.00

Hynek Hermansky

Title and Abstract not available

A time-domain binaural signal detection model and its application to binaural temporal phenomena

Jeroen Breebaart, Steven van de Par, and Armin Kohlrausch

Eindhoven University of Technology
&
Philips Research Labs, Eindhoven

This poster presents a time-domain binaural signal-detection model consisting of several subsequent signal-processingblocks that mimic the functional stages of the auditory system. The first block comprises a peripheral preprocessor, which contains, among other things, simple basilar membrane and inner haircell models. The second block compares the signals from the left and right ears by so-called Excitation-Inhibition (EI) units. These units have a very specific implementation to account for binaural sluggishness and interaural correlation sensitivity. The output of these units is corrupted by internal noise to limit the resolution. The third block analyzes the binaural and monaural inputs and extracts a decision variable from the input signals. For signal-detection purposes, information across frequency bands and time is combined according to an optimal criterion. It will be shown that the described model can account for the unification of binaural detection and lateralization data. Furthermore, it will be demonstrated that the choice of the implementation of the EI-units combined with the optimal detector is crucial to account for specific temporal detection phenomena. In particular, the model can account for teh apparent differences in the estimates of binaural time constants found with different experimental paradigms.

Dynamic loudness model for normal and hearing impaired listeners

Josef Chalupper* and Hugo Fastl

Institute for Human-Machine Communication, Technical University of Munich

Whereas dynamic loudness models for normal hearing listeners have existed for a long time, to date for hearing impaired listeners only models for predicting loudness of stationary sounds are available. Therefore, a dynamic loudness model (DLM) is proposed which predicts loudness of both stationary and non-stationary sounds for normal and hearing impaired listeners. In contrast to known stationary loudness models for hearing impaired listeners, in DLM only the specific loudness function is modified compared to normal hearing listeners. In order to fit the model to an individual hearing loss, standard audiometric data (threshold in quiet and categorical loudness scaling) are used. A comparison of the results of psychoacoustic experiments with model predictions shows that the proposed dynamic loudness model accounts for numerous aspects of loudness perception such as spectral and temporal integration of loudness and recruitment. These results for DLM suggest that it is possible also to model other hearing sensations such as sharpness, roughness and fluctuation strength for hearing impaired listeners.

*now with Siemens Audiological Engineering Group, Erlangen

Spectro-temporal processing of amplitude modulation

Stephan D. Ewert, Torsten Dau, Jesko L. Verhey, Birger Kollmeier

Arbeitsgruppe Medizinische Physik
Carl von Ossietzky Universität Oldenburg

The concept of a modulation filterbank following each peripheral filter was shown to qualitatively account for various empirical data in modulation detection. The present study addresses two basic questions that are still unclear. The first part examines whether the known frequency selectivity in the envelope domain reflects a ``true" spectral decomposition of the temporal envelope or whether it can be explained in terms of a temporal envelope-periodicity identifier (e.g. ACF) of limited resolution. It was tested whether a square-wave modulation with a temporal periodicity of 4 Hz leads to a higher threshold for a simultaneously presented sinusoidal 16-Hz target modulation than a 4-Hz sinusoidal masker modulation. The results are in line with a spectral decomposition of the temporal envelope, showing about 4 dB more masking produced by the square-wave modulation. The data can be explained by the harmonics of the square-wave modulation in the envelope-frequency domain. The second part is concerned with the role of the second-order envelope (referred to as the ``venelope" in the following) in modulation masking experiments. Venelope fluctuations can arise from either beats between the sinusoidal masker and target modulation, or from the intrinsic envelope fluctuations of a narrowband-noise modulator itself. It was tested to what extent the detection of venelope fluctuations is involved in corresponding tone-in-noise (TN), noise-in-tone (NT), and tone-in-tone (TT) modulation masking conditions. Thresholds in the NT condition are always lower than in the TN condition, comparable with the asymmetry of masking effect in the audio-frequency domain. In general, it was observed that (1) subjects use venelope fluctuations as a detection cue, and that (2) venelope fluctuations interfere with an additionally applied amplitude modulation when both fall in the same frequency range. To interpret the empirical findings, a general model structure for the processing of envelope and venelope fluctuations is proposed.

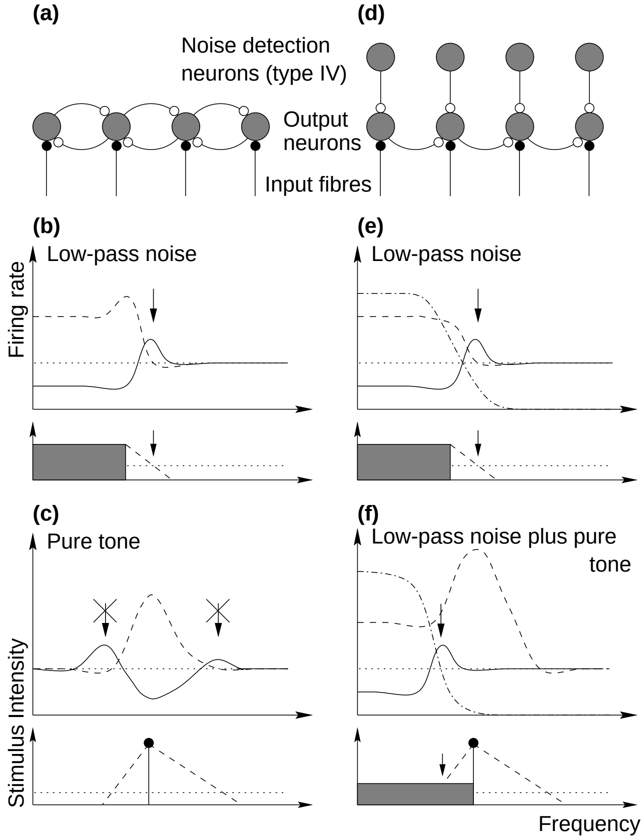
Zwicker Tone Illusion and Noise Reduction in the Auditory System

Jan-Moritz P. Franosch*, Richard Kempter*, Hugo Fastl**, J. Leo van Hemmen*

*Physik Department, TU München
D-85747 Garching, Germany
e-mail: jfranosc@ph.tum.de

**Lehrstuhl für Mensch-Maschine-Kommunikation, TU München
D-80333 München, Germany

The Zwicker tone is an auditory aftereffect. For instance, after switching off a broad-band noise with a spectral gap, one perceives it as a lingering pure tone, the pitch being in the gap [1, 2]. It is a unique illusion in that it cannot be explained by properties of the auditory periphery alone and has no direct analog in the visual system either. Here we present psychoacoustic experiments that reveal the crucial role of noise [3, 4]. Habituation is ruled out as a driving mechanism. Furthermore, we propose a neuronal model that predicts both the pitch and whether a sound can generate a Zwicker tone at all. We show that dominantly unilateral inhibition in conjunction with a neuronal noise-reduction mechanism explains the effect.



The figure shows both habituation (left column) and noise reduction (right).
(a) Neuronal implementation of simple (symmetric) lateral inhibition. Grey circles denote neurons, small filled circles indicate excitatory synapses and small open circles inhibitory synapses.
(b),(c) Response (upper panel) of the habituation model (a) to two sounds shown in the lower panels. Firing rates of the neurons before (horizontal dotted line, spontaneous rate), during (dashed line) and immediately after (solid line) the sound presentation are shown schematically. Downward arrows indicate Zwicker tones predicted by the model. In the case of the pure tone in (c) the habituation model predicts even two Zwicker tones (crossed arrows) whereas in reality there is none.
(d) Neuronal implementation of the full model with asymmetric inhibition and noise detection.
(e),(f) Response (upper panels) of the model in (d) to two sounds (lower panels). Dash-dotted lines indicate schematic firing rates of noise-detection neurons.

[1] Zwicker, E. 'Negative afterimage' in hearing. *J. Acoust. Soc. Am.* **36**, 2413–2415 (1964).
[2] Zwicker, E. & Fastl, H. *Psychoacoustics: Facts and Models.* Springer, Berlin (1999).
[3] Franosch, J.-M. P., Kempter, R. & van Hemmen, J. L. Zwicker tone illusion and noise reduction in the auditory system (2000). Preprint TU Mnchen.
[4] Fastl, H., Patsouras, D., Franosch, J.-M. P. & van Hemmen, J. L. Zwicker tones for pure tone plus band limited noise. In *Physiological and Psychophysical Bases of Auditory Function, Proc. of the 12th International Symposium on Hearing*, pp. 57–63 (2000).

Temporal integration in the European starling

A. Friebe, U. Langemann, and G.M. Klump

Lehrstuhl für Zoologie, TU München, Lichtenbergstr. 4, D-85748 Garching, Germany

The aim of this study was to describe the temporal integration in the European starling using the model of the ”temporal window” (Moore et al., 1988, J Acoust Soc Am 83: 1102-16). This model is the temporal analogue to the auditory filter in the frequency domain. The shape of the ”temporal window” determines how the auditory system weighs acoustic events presented at different time periods. On the basis of two experiments we investigated both the shape and width of this ”temporal window”.

In the first experiment we presented stimuli comparable to those used by Moore et al. (1998) to estimate the shape of the starling's ”temporal window”. Five starlings were trained in a Go/NoGo paradigm to report the detection of a 10 ms 2 kHz-signal (5 ms raised cosine ramps) that was presented at the beginning, the center or at the end of a temporal gap between two noise maskers (200 ms duration, 2 ms raised cosine ramps, 45 dB spectrum level). Thresholds for detecting the short probe signal were measured in relation to the duration of the gap and the temporal position of the probe. The probe signal was presented in 3-dB steps according to the method of constant stimuli. Signal detection theory was applied to determine detection thresholds; threshold criterion was a d' of 1.8 (which also applies to the second experiment). The starling’s detection thresholds were nearly identical when the signal was presented in the temporal center or at the beginning of the gap provided that the time interval between the end of the first noise pulse and the probe signal was the same. Thresholds decreased from about 52 dB to about 37 dB with increasing time interval from 10 ms to 50 ms, respectively. Thresholds were reduced by 3 to 10 dB for signals presented at the end of the gap. On the basis of these thresholds we calculated the shape of the temporal window as described by Moore et al. (1988). In the starling we obtained a ”temporal window” that is similar in shape but narrower than the human ”temporal window”. The equivalent rectangular duration (ERD) of the starling’s ”temporal window” is 5 ms compared to 8 ms in humans (Moore et al., 1998).

In the second experiment determining the critical masking interval, we used stimuli comparable to Penner and Cudahy (1973, J Acoust Soc Am 54: 1530-1534). The same 5 individual starlings were trained to detect a short click, presented in the temporal center of a wide band noise of different durations (0.4 to 300 ms; 40.1 dB spectrum level). Thresholds for detecting the click (presented in 3-dB steps, method of constant stimuli) were determined as a function of the duration of the noise. For noise durations from 30 to 300 ms thresholds increased with decreasing noise durations from 63 to 66 dB, respectively. This can be explained with an overshoot effect, that is known to show up at short noise durations. For noise durations below 30 ms thresholds decreased with decreasing noise durations from 65 to 60 dB. These noise durations are shorter than the duration of the ”temporal window” and therefore faciliate the detection of the click. Also the second experiment suggests that the temporal integration in the starling is slightly better than in humans.

(Supported by grants from the DFG: GRK 267, FG 306)

Reduction of dynamic cues in auralized binaural signals

Markus Fruhmann, Philip Mackensen**, Günther Theile*

Institut für Rundfunktechnik, München, Germany

A data-based auralization system can be realized by the convolution of binaural room impulse responses with the audio signal according to listener's head orientation. If head rotations are captured by means of headtracking, front-back inversions and in-head localisation are avoided compared to static auralization systems.
To save processing power it was investigated, whether the complete impulse responses must be calculated dynamically, or fading to a 'static' part after processing the first part dynamically is possible without producing audible artefacts. Experiments were done with measured binaural responses of a studio listening room. The length of the dynamic part of the binaural responses was varied to find the threshold for audible errors due to static calculation. A minimum length was found. Analysing the impulse response pattern, the minimum length of time for dynamic calculation coincides with the time of arrival of early reflections at listener's ears and thus the dimensions of the auralized room.

*now: AG Technische Akustik, MMK, TU München, e-mail: fruhmann@mmk.ei.tum.de

**now: T-Systems debis Systemhaus GEI GmbH, München

Influence of component phase on the loudness of complex tones

Hedwig Gockel, Brian C.J. Moore, Roy D. Patterson

Hedwig Gockel, Roy D. Patterson
CNBH Centre for the Neural Basis of Hearing Physiology Department, University of Cambridge, Downing Street, Cambridge, CB2 3EG, UK

Brian C.J. Moore
Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK

Sounds with strongly modulated envelopes undergo differential compression in the peripheral auditory system; medium to high level portions are amplified less than low level portions. One might expect this to affect the loudness of sounds with the same magnitude spectrum but differing phase spectra; phases giving waveforms with a high peak factor would be expected to lead to lower loudness than phases giving a lower peak factor. This was investigated by asking four listeners to match the loudness of complex tones and noise. The complex tones had a fundamental frequency of 62.5 Hz and were filtered into a frequency range from 625 Hz (10th harmonic) to 5000 Hz. The Gaussian noise was filtered in the same way. The components of the complex tones were added either in cosine phase (CP), giving a high peak factor, or in random phase (RP), giving a lower peak factor. Three tasks were used: (1) matching the loudness of the CP tones with that of the RP tones, and vice versa; (2) matching the loudness of the CP tones with that of the noise, and vice versa; (3) matching the loudness of the RP tones with that of the noise, and vice versa. The task were performed using six different levels of the fixed stimulus, ranging from 30 dB SPL to 80 dB SPL in 10 dB steps. Results were not consistent with the above hypothesis. CP tones were adjusted to a lower rms level than RP tones and than noise at the point of equal loudness, indicating that at equal levels the CP tones would sound louder. RP tones were adjusted to a lower level than noise at the point of equal loudness. The differences in level were greatest for mid-range levels of the fixed stimulus. In a second experiment, a continuous bandpass noise was present, in order to mask distortion products which might have caused the loudness differences. The bandpass masker reduced the observed differences in loudness somewhat, but the basic pattern of results was unchanged.

Loudness perception of fluctuating sounds

Giso Grimm, Jesko L. Verhey* and Volker Hohmann

Medizinische Physik , Universität Oldenburg
* CNBH, University of Cambridge, UK

Since most of our environmental sounds are not stationary and show temporal envelope fluctuations, it is important to know how theses temporal modulations affect psychoacoustical sensations like loudness. This topic has been intensively studied in the literature. However, the available data on the loudness of fluctuating sounds are not very consistent and even some contradictory results are reported, which might in part be due to methodological differences in the experimental design. Furthermore, it is not clear from the literature whether existing loudness models can account for the temporal effects in loudness perception. The aim of this study was therefore to enlighten some new aspects of time-dependent loudness using a uniform experimental framework. Loudness-matching data from normal-hearing subjects using amplitude modulated sounds and sounds with intrinsic envelope fluctuations will be presented. The influence of the crest factor and of the spectral content of the modulated sounds on the loudness were systematically quantified and compared with model predictions from recent loudness models.

Frequency analysis and synthesis using a Gammatone filterbank

Volker Hohmann

Medizinische Physik , Universität Oldenburg

Gammatone filters (Patterson et al., 1987) are widely used in computational auditory models for modeling the peripheral filtering in the Cochlea. This technical paper describes an efficient implementation of Gammatone filters, that is used in the peripheral filterbank stage of the Oldenburg perception model (PEMO). Additionally, a novel approach to the recombination of the Gammatone filterbank output is introduced that allows for the re-synthesis of the signal with a total time delay of 4ms. The signal reconstruction is nearly perfect, i.e., the difference between input and reconstructed output is merely audible. Possible application of the analysis/synthesis system introduced here is speech processing for hearing aids.

Literature:
Patterson, R. D., J. Nimmo-Smith, Holdsworth, J. and Rice, P. (1987). An efficient auditory filterbank based on the gammatone function. Technical report: paper presented at a meeting of the IOC Speech Group on Auditory Modelling at RSRE.

Human auditory brainstem potentials using optimized stimuli to compensate for basilar-membrane dispersion

Dirk Junius, Oliver Wegner, Torsten Dau

Arbeitsgruppe Medizinische Physik
Carl von Ossietzky Universität Oldenburg

It has recently been shown that a rising chirp stimulus, designed to activate synchronous discharges of auditory nerve fibers along the length of the human cochlear partition, evokes a significantly larger wave-V amplitude than a click presented at the same sensation level. The present study investigates potentials evoked by long-duration stimuli (up to about hundred ms duration) comprising this chirp. In one experiment, the stimulus consisted of a continuous alternating sequence of chirps: each rising chirp was followed by the temporally reversed (falling) chirp. In another experiment, a single rising chirp was temporally and spectrally embedded in a long-duration stimulus comprised of a 270 Hz tone, the 270-to-4000 Hz chirp, and a 4-kHz tone. In both of these experiments, the transitions between stimulus components were continuous. Results showed "peaked" response patterns whereby amplitude and (relative) latency of the peaks directly corresponded to those obtained with "single" chirps. These results clearly demonstrate the importance of considering the effects of the basilar-membrane traveling wave on the formation of the ABR, at least for wave V. At high stimulation levels, the response pattern from the second experiment also contained a periodic part preceding the peak, reflecting the frequency following response (FFR) of the 270-Hz tone. It is assumed that the same mechanisms are responsible for the generation of ABR-wave-V and FFR. A model for ABR/FFR generation is presented, which integrates synchronized neural activity across frequency after (level-dependent) basilar-membrane and auditory-nerve processing. The model accounts for the shape of the response pattern obtained in various stimulus conditions.

Comparison of Hearing Thresholds by Re-calibration of the Audiometric Equipment on an Individual Basis

B. L. Karlsen, M. Lydolf and A. O. Santillàn

Department of Acoustics, Aalborg University, Fredrik Bajers Vej 7 B4,
DK-9220 Aalborg Ø, Denmark, Email: blyk@acoustics.auc.dk

Currently, hearing thresholds are measured with different earphones in the hearing clinics leading to different thresholds. In this paper, five different audiometric earphones are compared by using threshold measurements. The five earphones used in the experiment were calibrated on a coupler conforming to IEC 60318-1 (1998). The frequency response of the earphones on the coupler were also recorded. Furthermore, the phone-to-closed-ear-canal-

transfer-function (PTF) of all earphones on all subjects was measured. In order to re-calibrate the audiometric equipment to the closed ear canal, the mean PTF across subjects is added and the coupler response is subtracted from the mean absolute threshold data. The resulting absolute threshold data for the five earphones are much more similar than they were before the re-calibration.

Capturing spectro-temporal modulations for automatic speech recognition

Michael Kleinschmidt

Arbeitsgruppe Medizinische Physik
Carl von Ossietzky Universität Oldenburg

Psychoacoustical and neurophysiological results indicate that spectro-temporal modulations play an important role in sound perception. Speech signals, in particular, exhibit distinct spectro-temporal patterns which are well matched by receptive fields of cortical neurons (A1).

In order to improve the performance of automatic speech recognition (ASR) systems a number of different approaches are presented, all of which target at capturing spectro-temporal modulations. By deriving secondary features from the output of a perception model the tuning of neurons towards different envelope fluctuations is modeled.

The following types of secondary features are introduced: product of two or more windows of variable size (sigma-pi cells) in the spectro-temporal representation, fuzzy-logical combination of windows and also two-dimensional Gabor function filtering.

The different approaches are introduced and tested on a simple isolated word recognition task. Using an iterative optimisation scheme allows for analysis of secondary feature parameters most suitable for the given problem.

Parameter Opimization of the Low-level Acoustic Reflex Measurement

Matthias Müller-Wehlau, Manfred Mauermann, Birger Kollmeier

Arbeitsgruppe Medizinische Physik
Carl von Ossietzky Universität Oldenburg

In 1996 J. Neumann et al. described a new method to detect the acoustic reflex[1] by means of techniques usually employed for recording of otoacoustics emissions using two identical short stimuli to elicit and detect the AR rather than using two different tones as done by the commonly utilized method. In this new method the two identical stimuli are presented in a short succession. Since the change in impedance due to the acoustic reflex has a latency of some tens of milliseconds the second tone pulse is presented after a sufficient long time following the first thus leading to a maximal difference in the measured middle ear response between these two pulses. Great advantages towards the conventional measurement technique might be given by the possibility to calculate the mean of a number of succsessive representations thus increasing the signal noise ratio and to determine the change in phase between stimulus and probe tone.

The aim of this paper is to deliver an optimized set of parameters for this method and to provide different solutions to evaluate the yielded data to determine the acoustic reflex threshold. For this nummerous measurements have been conducted with several subjects using different setups. The subjects, aged from 21 to 34 years, had normal pure-tone hearing threshold.

To process the data a variety of algorithm are presented, different in terms of the used artefact suppression. The ART is determined by using two alternative criteria. The data gained with these methods are compared among each other and with the conventionaly used measurement showing the advantages of the new method.

[1] **Neumann, J. , Uppenkamp, S., Kollmeier, B.:** Detection of the acoustic reflex below 80dB HL, Audiology and Neurootology 1, 359-369, 1997

Threshold Methods For TTS Evaluations

Rodrigo Ordoñez, Brian Lykkegaard Karlsen, Morten Lydolf and Dorte Hammershøi

Department of Acoustics, Aalborg University. Fredrik Bajers Vej 7 B4
DK-9220 Aalborg Ø Denmark

To measure in a reliable manner a small amount of Temporary Threshold Shift (TTS), the audiometric method used must be fast, accurate and repeatable. This is because of the time varying characteristics of the TTS and because only small amounts of TTS can be induced in subjects under controlled and safe conditions. To assess the influence of the audiometric stimuli on the standard error of the mean of the determined threshold, the \textit{Bekesy} audiometric method will be used with three different stimuli. Frequency modulated tones, band-passed noise and pure tones will be compared. The three types of stimuli will also be constructed as single events (< than 20 ms), bursts (a train of single events) and normal stimuli (1 sec.). Threshold determinations will be carried out for octave bands between 0.5 and \$8~kHz\$. In a preliminary experiment a few subjects will be used to achieve a high number of repetitions.

A comparison of experimental methods for measuring auditory-visual (a)synchrony perception

Steven van de Par, Armin Kohlrausch, and James F. Juola

Eindhoven University of Technology
&
Philips Research Labs, Eindhoven

Experiments will be presented that are concerned with the perception of auditory-visual (AV) asynchrony. Subjects were presented with AV stimuli which consisted of a disk on a screen that accelerated downwards towards a bar, and at the moment of incidence rebounded upwards while decelerating. Simultaneously, a short tone was presented with a sharp onset and an exponentially decaying envelope. The relative timing of the tonal onset with respect to the moment of visual incidence was varied in 15 steps between video leading by 350 ms and audio leading by 350 ms. A number of experimental methods were compared which have been used previously to establish the point of subjective equality (PSE), which is defined as that physical delay between audio and video that leads to the percept of synchrony. In all experiments, the above AV stimulus was used in a single interval procedure. The experiments differed in the response categories that the subject could use. In the first experiment, three categories were offered: audio first, video first, or synchronous (AVS). In the second experiment responses were either synchronous or asynchronous (SA), and in the last experiment two answers identifying the temporal order (TO) could be given, namely audio first or video first. Careful study of the data suggest that subjects have a consistent strategy for responding within the first two experiments (AVS and SA) and that estimates of the PSE agreed well for these two procedures. For the last experiment (TO), however, results suggest that subjects can adopt different response strategies which lead to widely varying estimates for the PSE. Based on these results, we conclude that this latter method is less suitable for studying AV temporal perception.

Lokalisation Experiments with a six-microphone array to simulata dummy head perfomance.

Natalia Podlaszewski, Volker Mellert

WG Acoustic, Dept of Physics, Univ. of Oldenburg, 26111Oldenburg, Germany

Ear-related sound recordings are used to reproduce the spatial information of a given sound field. The recordings are usually performed with the help of a dummy head, which is matched in its acoustic properties to an average human listener. An new array of microphones substituting the dummy head allows to adjust the HRTFs to an individual listener, i.e. to the acoustic features of a single user. Localisation experiments with an optimized six-microphone array are reported. The localisation is measured with an adaptive procedure in the virtual acoustic space.

Auditory brain stem responses evoked by lateralized click stimuli

Helmut Riedel und Birger Kollmeier

AG Medizinische Physik Carl von Ossietzky Universität Oldenburg
D-26111 Oldenburg

The dependence of binaurally evoked auditory brain stem responses and the binaural difference potential on simultaneously presented interaural time- and level differences is investigated in order to assess the representation of stimulus lateralization in the brain stem. Auditory brain stem responses to binaural click stimuli with all combinations of three

interaural time- and three interaural level differences were recorded from 12 subjects and 4 channels. The latency of Jewett wave V is smallest for zero interaural time difference and largest for the trading stimuli. The amplitude of wave V is largest for centrally perceived stimuli, i.e., the diotic and trading stimuli, and smallest for the most laterally perceived stimuli. The latency of the most prominent peak of the binaural difference potential DN1 mainly depends on the interaural time difference. The amplitude of the components of the binaural difference potential, DP1-DN1, depends similarly on stimulus conditions as wave V amplitude in the case of the binaural stimuli: smallest amplitudes are found for the most lateral stimuli and largest amplitudes for central stimuli. The dependence of auditory brain stem reponses and the binaural difference potential on the lateralization of the stimuli shows that interaural level- and time differences are not processed independently. This supports the hypothesis that directional information is already extracted and represented at the level of the brain stem.

Effect of peripheral compression on middle latency auditory evoked fields (AEF)

Andre Rupp
Sektion Biomagnetismus, Neurologische Klinik
der Universität Heidelberg

and

Volker Hohmann
Medizinische Physik , Universität Oldenburg

The aim of this project is to measure the influence of peripheral compression on on-frequency and off-frequency forward masking of high-frequency tone bursts using middle latency AEF as recorded by whole-head magnetoencephalography (MEG).

Oxenham and Plack (1997) showed psychoacoustically that on-frequency maskers lead to a steeper increase of masker level as a function of test tone level at masked threshold than off-frequency maskers. This finding is consistent with the assumption that peripheral compression is mainly active for frequencies in the region of CF but not for frequencies well below CF. This asymmetry vanishes in hearing-impaired listeners, which is most likely due to a loss in peripheral compression, leading to a linearization of the system.

We adapted the stimuli used by Oxenham and Plack so that they are suitable for MEG measurements. A 3.8kHz tone burst is used as test tone and lowpass noises with 2.5kHz and 4kHz cut-off frequency are used as maskers for the off-frequency and on-frequency conditions, respectively. Psychoacoustical data are presented that show the asymmetry of masking for the two conditions using the adapted stimuli, indicating that they are in principle suitable for revealing peripheral compression. Preliminary MEG results from a pilot study with normal-hearing subjects showed that the test tone evokes middle latency AEF’s and that dipole moment magnitude varies with the test tone level. The dependence of the AEF-magnitude on test tone level will be compared for the two masking conditions. Further studies will include MEG measurements with hearing-impaired subjects that aim at assessing the loss of peripheral compression objectively using AEF's.

Literature:

Oxenham, A.J. and Plack, C.J. (1997). A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing. J. Acoust. Soc. Am. 101 (6), 3666-3675.

A new method for localization studies and its application to audiology

Bernhard Seeber

AG Technische Akustik, MMK, Technische Universität München, email: seeber@ei.tum.de

Many scientific studies and technical applications investigate the acoustical localization in the field of vision. Therefore it is suitable to display the perceived auditory direction by a light point. In the formerly known methods subjects point with a light pointer directly out of their hands or with a pointer mounted on a revolvable axis in front of them. Using these methods influences of the motor system of the subject and the parallax are possible. The calibration of the system and the recording of the results also turn out to be difficult. The here proposed new method utilizes a laser pointer with a deflection unit instead, which is controlled by a computer. Subjects enter their perceived direction with a trackball. The laser spot moves according to the turning of the ball smoothly on a defined track. A complicated mechanical calibration can be avoided by calibrating the deflection unit by a computer and the results are ready for evaluation. The intuitive experimental operation and the high resolution of the data makes this method particularly suitable for localization research in audiology. Selected localization results from eleven patients with cochlear implant and conventional hearing aid on the contralateral side are presented. It was found that some of these patients are able to integrate the signals from both sides to a perception of sound direction. One of the subjects was even able to localize with an accuracy close to that of normal hearing persons.

Modulation masking produced by masker modulation tone complexes

Jesko L. Verhey*), Stephan Ewert, and Torsten Dau

Arbeitsgruppe Medizinische Physik, Carl von Ossietzky
Universität Oldenburg, Postfach 2503, D-26111 Oldenburg, Germany
*) Centre for the Neural Basis of Hearing, The Physiological Laboratory, Cambridge, CB2 3EG, United Kingdom

Thresholds were measured for detecting a sinusoidal amplitude modulation in the presence of a masker-modulation tone-complex. Both modulations were applied to the same sinusoidal carrier. Two different masker modulations were used: i) a pair of components beating at the signal-modulation frequency and ii) a three-tone complex producing a sinusoidal amplitude modulation in the modulation domain whereby the frequency of the slow amplitude variation was equal to the signal-modulation frequency. Masked thresholds were measured as function of the phase of the signal modulation relative to the slow amplitude variation in the masker modulation. In all conditions, thresholds were lower for an in-phase signal modulation compared to an anti-phase condition. The maximum threshold difference was 15 dB. The possible role of nonlinearities prior to a modulation filterbank are discussed in the context of the present findings.

Influence of interfering noise and measurement procedure on speech intelligibility

Kirsten Wagener, Thomas Brand, Birger Kollmeier

AG Medizinische Physik Universität Oldenburg

In order to investigate the comparability of speech tests the influence of non speech parameters on speech intelligibility was tested using the Oldenburg sentence test (Wagener et al., 2000, Contributions to Psychological Acoustics - 8th Oldenburg Symposium on Psychological Acoustics). Measurements with 10 normal hearing and 10 hearing impaired subjects were performed. The speech reception threshold (SRT: signal to noise ratio that yields 50% intelligibility) and the respective slope at the SRT were investigated. The parameters absolute noise level, interfering noise (non modulated and amplitude modulated speech simulating noises), and adaptive procedure (constant noise level versus constant speech level, continuous noise versus interrupted noise synchronous to the sentence pauses) were varied.

As expected, there is no effect of the absolute level except for absolute levels near the hearing threshold. The strongly modulated noise yields a 14 dB lower SRT than the non modulated noise for normal hearing subjects. Some of the hearing impaired subjects show a comparable profit using modulated noise, but there are some subjects who did not benefit from the modulations. To investigate this splitting up more precisely, more experiments were performed with the hearing impaired subjects using modified amplitude modulated noise (with reduced pauses) and measurements in quiet. The SRT results using strongly amplitude modulated noise are highly correlated with the SRT results in quiet. However, the variability across hearing impaired listeners with regard to their performance in fluctuating noise can not completely be explained by this simple model.

Comodulation masking release in Mongolian gerbils (*Meriones unguiculatus*) studied with narrow-band maskers

E. Wagner, G.M. Klump

Lehrstuhl für Zoologie, TU München, Lichtenbergstr.4, D-85748 Garching, Germany

Comodulation masking release (CMR) is a mechanism developed in the evolution of the hearing system to improve signal detection in interfering background noise. CMR describes the reduced masking of a pure tone when the components of a masker show the same amplitude modulation in different frequency regions.

Four Mongolian gerbils (*Meriones unguiculatus*, age during the time of testing 33-39 month) were trained in an operant Go/NoGo procedure with food rewards to report a pure-tone stimulus of 2 kHz (410 ms total duration, 10 ms raised cosine ramps) in a continuous masking noise (spectrum level 40 dB/Hz). The masker consisted of two narrow-band noise stimuli (bandwidth 25 Hz). One noise band was centered on the signal frequency (on-frequency band). The center frequency of the other noise band (flanking band) was 400, 1200, 1600, 1800, 1900, 2100, 2200, 2400, 2800 or 3600 Hz. Uncorrelated maskers were created by time shifting the envelope of the flanking band in relation to the envelope of the on-frequency band by 2000 ms. Test stimuli were presented according to the method of constant stimuli. The thresholds were determined using signal-detection theory and a d' of 1,8 as the threshold criterion. The amount of masking release was defined as the threshold difference between the correlated and uncorrelated condition.

In this flanking-band paradigm gerbils show a considerable CMR over at least 4 critical bandwidths (the gerbils critical band at 2kHz is 218 Hz, see Kittel 2000, Diplomthesis, TU München). The observed pattern of CMR in the gerbil is similar to that found in the starling (Klump et al. 2001, in: Houtsma et al.: Physiological and psychophysical bases of auditory function) and in the human (e.g. Schooneveldt & Moore 1987, J Acoust Soc Am 82: 1944-1956), the amount of CMR is even bigger.

(Supported by grants from the DFG: GRK 267, FG 306. We thank Ingo Stürmer and Joan M. Sinnott for providing gerbils for this study that are offspring of gerbils collected in the field and only bred for up to 4 generations in captivity.)

A model for the generation of middle latency and steady-state auditory evoked potentials

Oliver Wegner, Torsten Dau

AG Medizinische Physik, Carl von Ossietzky Universität Oldenburg

This study presents a model for the generation of auditory evoked middle latency responses (MLR). The model is an extension of the model for ABR generation by Dau (2001). It is based upon the concept introduced by Goldstein and Kiang (1958) that evoked potentials recorded at remote electrodes can theoretically be given by convolution of an elementary unit waveform (unitary response) with the instantaneous discharge rate function for the corresponding unit. The unitary response - representing activity from the primary auditory cortex - was computed once by deconvolving experimental click-evoked MLR data with the simulated click-evoked compound auditory-nerve discharge rate function. MLR were predicted by convolving the (stimulus-dependent) discharge rate functions, calculated with the model by Heinz et al. (2001), with the above unitary response. This was done for a number of stimulus configurations including chirps and clicks presented at different repetition rates as well as AM tones of different modulation rates. The results demonstrate the importance of peripheral signal processing for the formation of MLR.

Dzhafarov-Colonius theory of Fechnerian scaling: Empirical validity of axioms in a unidimensional stimulus domain - auditory intensity discrimination.

Karin Zimmer & Hans Colonius

Institut fuer Kognitionsforschung
Carl von Ossietzky - Universitaet Oldenburg

Recently, Dzhafarov and Colonius [Psych. Bull. & Rev.,6, 239-268 (1999)] proposed a theory deriving the perceptual distance of stimuli from discriminability measures in both uni-, and multi-dimensional stimulus domains. In the theory, the existence of a Fechnerian metric rests on assumptions about the shape of the psychometric functions which are empirically testable. Thus, "same-different" psychometric functions (1) have to attain a single minimum at some point, (2) must have the same psychometric order, i.e. must show the same degree of cuspidness, or flatness, along the stimulus continuum, and (3) should be odd-symmetric around their minima. In the present investigation, these assumptions were evaluated experimentally employing acoustical stimuli differing in one stimulus dimension, intensity. In a "same-different" paradigm, using the method of constant stimuli, seven 1000 Hz-tones deviating from 0 dB to 2.1 dB with respect to a standard were presented diotically to five observers. For every participant, psychometric functions were assessed at five hearing levels ranging from 40 dB HL to 80 dB HL. Results support the assumptions of the theory in four of five observers. Thus, data such as these are in principle suitable for deriving a loudness scale from discriminability measures.