Adaptive Processes in Hearing

Our auditory environment is constantly changing and evolving over time, requiring us to rapidly adapt to a complex dynamic sensory input. This adaptive ability of our auditory system can be observed at different levels, from individual cell responses to complex neural mechanisms and behavior, and is essential to achieve successful speech communication, correct orientation in our full environment, and eventually survival. These adaptive processes may differ in individuals with hearing loss, whose auditory system may cope via “readapting” itself over a longer time scale to the changes in sensory input induced by hearing impairment and the compensation provided by hearing devices. These devices themselves are now able to adapt to the listener's individual environment, attentional state, and behavior. These topics related to auditory adaptation, in the broad sense of the term, were central to the 6th International Symposium on Auditory and Audiological Research held in Nyborg, Denmark, in August 2017. The symposium addressed adaptive processes in hearing from different angles, together with a wide variety of other auditory and audiological topics. The papers in this special issue result from some of the contributions presented at the symposium.
Cortical oscillations and entrainment in speech processing during working memory load

Neuronal oscillations are thought to play an important role in working memory (WM) and speech processing. Listening to speech in real-life situations is often cognitively demanding but it is unknown whether WM load influences how auditory cortical activity synchronizes to speech features. Here, we developed an auditory n-back paradigm to investigate cortical entrainment to speech envelope fluctuations under different degrees of WM load. We measured the electroencephalogram, pupil dilations and behavioural performance from 22 subjects listening to continuous speech with an embedded n-back task. The speech stimuli consisted of long spoken number sequences created to match natural speech in terms of sentence intonation, syllabic rate and phonetic content. To burden different WM functions during speech processing, listeners performed an n-back task on the speech sequences in different levels of background noise. Increasing WM load at higher n-back levels was associated with a decrease in posterior alpha power as well as increased pupil dilations. Frontal theta power increased at the start of the trial and increased additionally with higher n-back level. The observed alpha-theta power changes are consistent with visual n-back paradigms suggesting general oscillatory correlates of WM processing load. Speech entrainment was measured as a linear mapping between the envelope of the speech signal and low-frequency cortical activity (}
Decoding the auditory brain with canonical component analysis

The relation between a stimulus and the evoked brain response can shed light on perceptual processes within the brain. Signals derived from this relation can also be harnessed to control external devices for Brain Computer Interface (BCI) applications. While the classic event-related potential (ERP) is appropriate for isolated stimuli, more sophisticated "decoding" strategies are needed to address continuous stimuli such as speech, music or environmental sounds. Here we describe an approach based on Canonical Correlation Analysis (CCA) that finds the optimal transform to apply to both the stimulus and the response to reveal correlations between the two. Compared to prior methods based on forward or backward models for stimulus-response mapping, CCA finds significantly higher correlation scores, thus providing increased sensitivity to relatively small effects, and supports classifier schemes that yield higher classification scores. CCA strips the brain response of variance unrelated to the stimulus, and the stimulus representation of variance that does not affect the response, and thus improves observations of the relation between stimulus and response.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Google AI for Perception, United States., Ecole Normale Superieure, University of Rochester
Authors: de Cheveigné, A. (Ekstern), Wong, D. D. E. (Ekstern), Di Liberto, G. M. (Ekstern), Hjortkjær, J. (Intern), Slaney, M. (Ekstern), Lalor, E. (Ekstern)
Pages: 206-216
Publication date: 2018
Main Research Area: Technical/natural sciences

Publication information
Journal: Neuroimage
Volume: 172
ISSN (Print): 1053-8119
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 6.31 SJR 3.823 SNIP 1.752
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 4.48 SNIP 1.84 CiteScore 6.71
Effects of Hearing Loss and Fast-Acting Compression on Amplitude Modulation Perception and Speech Intelligibility

Objective: The purpose was to investigate the effects of hearing-loss and fast-acting compression on speech intelligibility and two measures of temporal modulation sensitivity.

Design: Twelve adults with normal hearing (NH) and 16 adults with mild to moderately severe sensorineural hearing loss were tested. Amplitude
modulation detection and modulation-depth discrimination (MDD) thresholds with sinusoidal carriers of 1 or 5kHz and modulators in the range from 8 to 256 Hz were used as measures of temporal modulation sensitivity. Speech intelligibility was assessed by obtaining speech reception thresholds in stationary and fluctuating background noise. All thresholds were obtained with and without compression (using a fixed compression ratio of 2:1).

Results: For modulation detection, the thresholds were similar or lower for the group with hearing loss than for the group with NH. In contrast, the MDD thresholds were higher for the group with hearing loss than for the group with NH. Fast-acting compression increased the modulation detection thresholds, while no effect of compression on the MDD thresholds was observed. The speech reception thresholds obtained in stationary noise were slightly increased in the compression condition relative to the linear processing condition, whereas no difference in the speech reception thresholds obtained in fluctuating noise was observed. For the group with NH, individual differences in the MDD thresholds could account for 72% of the variability in the speech reception thresholds obtained in stationary noise, whereas the correlation was insignificant for the hearing-loss group.

Conclusions: Fast-acting compression can restore modulation detection thresholds for listeners with hearing loss to the values observed for listeners with NH. Despite this normalization of the modulation detection thresholds, compression does not seem to provide a benefit for speech intelligibility. Furthermore, fast-acting compression may not be able to restore MDD thresholds to the values observed for listeners with NH, suggesting that the two measures of amplitude modulation sensitivity represent different aspects of temporal processing. For listeners with NH, the ability to discriminate modulation depth was highly correlated with speech intelligibility in stationary noise.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Widex A/S
Authors: Wiinberg, A. (Intern), Jepsen, M. L. (Ekstern), Epp, B. (Intern), Dau, T. (Intern)
Number of pages: 10
Publication date: 2018
Main Research Area: Technical/natural sciences

Publication information
Journal: Ear and Hearing
ISSN (Print): 0196-0202
Ratings:
  BFI (2018): BFI-level 2
  Web of Science (2018): Indexed yes
  BFI (2017): BFI-level 2
  Web of Science (2017): Indexed Yes
  BFI (2016): BFI-level 2
  Scopus rating (2016): CiteScore 2.97 SJR 1.865 SNIP 1.571
  Web of Science (2016): Indexed yes
  BFI (2015): BFI-level 2
  Scopus rating (2015): SJR 1.753 SNIP 2.008 CiteScore 2.94
  BFI (2014): BFI-level 2
  Scopus rating (2014): SJR 1.93 SNIP 1.726 CiteScore 2.86
  Web of Science (2014): Indexed yes
  BFI (2013): BFI-level 2
  Scopus rating (2013): SJR 1.893 SNIP 2.109 CiteScore 3.18
  ISI indexed (2013): ISI indexed yes
  BFI (2012): BFI-level 2
  Scopus rating (2012): SJR 1.918 SNIP 1.708 CiteScore 2.95
  ISI indexed (2012): ISI indexed yes
  Web of Science (2012): Indexed yes
Effects of musical training and hearing loss on pitch discrimination

Our ability to perceive the pitch of complex sounds is essential for melody perception and for our enjoyment of music. It also plays an important role in speech perception to convey intonation and sometimes meaning, e.g., in tonal languages, and greatly helps segregation of competing sound sources. Humans are able to discriminate very small changes in the pitch of complex harmonic sounds, with fundamental frequency difference limens (F0DLs) that can be smaller than 1% of the fundamental frequency (F0). However, performance in such pitch discrimination tasks is known to depend on the harmonic content of the sound and whether the harmonics are resolved by the auditory frequency analysis operated by cochlear processing. F0DLs are also heavily influenced by the amount of musical training received by the listener and by the spectrotemporal auditory processing deficits that often accompany sensorineural hearing loss. This paper reviews the latest evidence for how musical training and hearing loss affect pitch discrimination performance, based on behavioral F0DL experiments with complex tones containing either resolved or unresolved harmonics, carried out in listeners with different degrees of hearing loss and musicianship. A better understanding of the interaction between these two factors is crucial to determine whether auditory training based on musical tasks or targeted towards specific auditory cues may be useful to hearing-impaired patients undergoing hearing rehabilitation.
cochlear synaptopathy after participating in a concert. Young adult listeners with hearing thresholds ≤ 20 dB HL between
0.25-8 kHz were recruited
and divided into two groups: listeners voluntarily participating in concerts and control listeners with no concert participation
during the study. Exposure was assessed with dosimeters in both groups for one event duration. Concert participants
were advised to use hearing protectors and exposure levels were determined from actual use. Listeners performed three
sessions of audiometry, auditory brainstem response (ABR), and speech in noise measurements. Session 1 was
performed within a week up to the concert, session 2 within 24 hours after the concert, and session 3 approximately 4
weeks after the
concert. We hypothesized that concert participants would show reduced level-growth of ABR wave I and that wave-I level-
growth would be a predictor of speech discrimination score. The data indicate that postexposure wave-I level-growth was
not reduced compared to pre-exposure values, and that neither were speech scores. Therefore, the results might suggest
that: a) concert goers do not develop cochlear synaptopathy in response to typical exposure from one event, or b)
synaptopathy occurs only for more severe exposure in humans, or c) the utilized measures are not sensitive enough to
detect the damage.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Acoustic Technology, Technical University of
Denmark
Authors: Smits, B. (Ekstern), Holtegaard, P. (Intern), Jeong, C. (Intern), Santurette, S. (Intern)
Number of pages: 6
Publication date: 2018

Frequency modulation excursion and rate discrimination in normal-hearing and hearing-impaired listeners
Most natural sounds contain frequency fluctuations over time such as changes in their fundamental frequency, non-
periodic speech formant transitions, or periodic fluctuations like musical vibrato. These are sometimes characterized as
frequency modulation (FM) with a given excursion (FMe) and rate (FMr) (Fig.1). Accurate
processing of FM may play an important role in music and speech perception, especially in complex instrument or talker
situations. While age and sensorineural hearing loss (SNHL) can affect FM detection thresholds [1,2] and SNHL can affect
the range of FMe and FMr values producing a sung vowel percept (Fig.2) [3], less is known about how these factors affect
FMe and FMr discrimination. Moreover, reference data for FM discrimination in normal-hearing (NH) listeners remains
scarce [4-6]. As discrimination tasks are closer to what listeners may use in real-life situations, this study investigated the
effects of age and SNHL on FMe and FMr difference limens (DLs) for reference values typical of frequency fluctuations
observed in speech and music signals.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Copenhagen University Hospital, Eriksholm
Research Centre
Authors: Schindwolf, I. (Ekstern), Vatti, M. (Ekstern), Santurette, S. (Intern)
Number of pages: 1
Publication date: 2018
Event: Poster session presented at 41st MidWinter Meeting of the Association for Research in Otolaryngology, San Diego,
United States.
Main Research Area: Technical/natural sciences
Electronic versions:
SchindwolfVattiSanturette_ARO2018.pdf

Bibliographical note
Published abstract and presented poster at 41st MidWinter Meeting of the Association for Research in Otolaryngology,
February 9-14, 2018, San Diego, CA — Abstract #PS 375.
Source: PublicationPreSubmission
Source-ID: 144368855
Publication: Research › Poster – Annual report year: 2018

Periphony-Lattice Mixed-Order Ambisonic Scheme for Spherical Microphone Arrays
Most methods for sound field reconstruction and spherical beamforming with spherical microphone arrays are
mathematically based on the spherical harmonics expansion. In many cases, this expansion is truncated at a certain order
as in higher order ambisonics (HOA). This truncation leads to performance that is independent of the incident direction of the sound waves. On the other hand, mixed-order ambisonic (MOA) schemes that select an appropriate subset of spherical harmonics can improve the performance for horizontal directions at the expense of other directions. This paper proposes an MOA scheme called Periphony-Lattice to improve sound field reconstruction performance for horizontally incident sound waves. The proposed scheme is compared with the previously introduced MOA and HOA schemes in terms of theoretical truncation error and performance in sound field reconstruction and spherical beamforming. Computer simulations and measurements are conducted with a spherical array of 52 microphones with a nonuniform layout. The results show that the proposed MOA scheme has better performance in sound field reconstruction and spherical beamforming for horizontal sound waves than the other schemes for a given number of microphones. This scheme can be applied to other spherical array layouts if the number of microphones is greater than that of the required spherical harmonics coefficients, and may improve the horizontal performance.
Robust speech dereverberation with a neural network-based post-filter that exploits multi-conditional training of binaural cues

This study presents an algorithm for binaural speech dereverberation based on the supervised learning of short-term binaural cues. The proposed system combined a delay-and-sum beamformer (DSB) with a neural network-based post-filter that attenuated reverberant components in individual time-frequency (T-F) units. A multi-conditional training (MCT) procedure was used to simulate the uncertainties of short-term binaural cues in response to room reverberation by mixing the direct part of head related impulse responses (HRIRs) with diffuse noise. Despite being trained with only anechoic HRIRs, the proposed dereverberation algorithm was tested in a variety of reverberant environments and achieved considerable improvements relative to a coherence-based approach in terms of three objective metrics reflecting speech quality and speech intelligibility. Moreover, a systematic evaluation showed that the proposed system generalized very well to a wide range of acoustic conditions, including various measured binaural room impulse responses (BRIRs) reflecting different reverberation times, azimuth positions spanning the entire frontal hemifield, various source-receiver distances as well as different artificial heads.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: May, T. (Intern)
Pages: 406-414
Publication date: 2018
Main Research Area: Technical/natural sciences

Publication information
Journal: IEEE/ACM Transactions on Audio, Speech, and Language Processing
Volume: 26
Issue number: 2
ISSN (Print): 2329-9290
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
Scopus rating (2016): CiteScore 3.5 SJR 0.887 SNIP 3.643
Scopus rating (2015): SJR 1.462 SNIP 2.437 CiteScore 2.4
Web of Science (2015): Indexed yes
Scopus rating (2014): SJR 1.437 SNIP 2.916 CiteScore 4.58
Scopus rating (2013): SJR 1.127 SNIP 3.123 CiteScore 4.03
ISI indexed (2013): ISI indexed no
Web of Science (2013): Indexed yes
Scopus rating (2012): SJR 0.902 SNIP 2.455 CiteScore 3.05
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
Scopus rating (2011): SJR 0.952 SNIP 2.735 CiteScore 2.56
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
Scopus rating (2010): SJR 1.34 SNIP 2.614
Scopus rating (2009): SJR 1.4 SNIP 2.981
Scopus rating (2008): SJR 1.459 SNIP 2.354
Scopus rating (2007): SJR 1.226 SNIP 2.711
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 1.256 SNIP 3.371
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 1.3 SNIP 3.667
Scopus rating (2004): SJR 0.982 SNIP 2.99
Task dialog by native-Danish talkers in Danish and English in both quiet and noise

The zip files contain recorded conversations between 19 pairs of normal-hearing native-Danish talkers taking part in an experiment in the lab of the Hearing Systems Group at The Technical University of Denmark during October-November 2016.

Task-Modulated Cortical Representations of Natural Sound Source Categories

In everyday sound environments, we recognize sound sources and events by attending to relevant aspects of an acoustic input. Evidence about the cortical mechanisms involved in extracting relevant category information from natural sounds is, however, limited to speech. Here, we used functional MRI to measure cortical response patterns while human listeners categorized real-world sounds created by objects of different solid materials (glass, metal, wood) manipulated by different sound-producing actions (striking, rattling, dropping). In different sessions, subjects had to identify either material or action categories in the same sound stimuli. The sound-producing action and the material of the sound source could be decoded from multivoxel activity patterns in auditory cortex, including Heschl’s gyrus and planum temporale. Importantly, decoding success depended on task relevance and category discriminability. Action categories were more accurately decoded in auditory cortex when subjects identified action information. Conversely, the material of the same sound sources was decoded with higher accuracy in the inferior frontal cortex during material identification. Representational similarity analyses indicated that both early and higher-order auditory cortex selectively enhanced spectrotemporal features relevant to the target category. Together, the results indicate a cortical selection mechanism that favors task-relevant information in the processing of nonvocal sound categories.
The impact of exploiting spectro-temporal context in computational speech segregation

Computational speech segregation aims to automatically segregate speech from interfering noise, often by employing ideal binary mask estimation. Several studies have tried to exploit contextual information in speech to improve mask estimation accuracy by using two frequently-used strategies that (1) incorporate delta features and (2) employ support vector machine (SVM) based integration. In this study, two experiments were conducted. In Experiment I, the impact of exploiting spectro-temporal context using these strategies was investigated in stationary and six-talker noise. In Experiment II, the delta features were explored in detail and tested in a setup that considered novel noise segments of the six-talker noise. Computing delta features led to higher intelligibility than employing SVM based integration and intelligibility increased with the amount of spectral information exploited via the delta features. The system did not, however, generalize well to novel segments of this noise type. Measured intelligibility was subsequently compared to extended short-term objective intelligibility, hit–false alarm rate, and the amount of mask clustering. None of these objective measures alone could account for measured intelligibility. The findings may have implications for the design of speech segregation systems, and for the selection of a cost function that correlates with intelligibility.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Bentsen, T. (Intern), Kressner, A. A. (Intern), Dau, T. (Intern), May, T. (Intern)
Pages: 248-259
Publication date: 2018
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of the Acoustical Society of America
Volume: 143
Issue number: 1
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
The Role of Place Cues in Voluntary Stream Segregation for Cochlear Implant Users

Sequential stream segregation by cochlear implant (CI) listeners was investigated using a temporal delay detection task composed of a sequence of regularly presented bursts of pulses on a single electrode (B) interleaved with an irregular sequence (A) presented on a different electrode. In half of the trials, a delay was added to the last burst of the regular B sequence, and the listeners were asked to detect this delay. As a jitter was added to the period between consecutive A bursts, time judgments between the A and B sequences provided an unreliable cue to perform the task. Thus, the segregation of the A and B sequences should improve performance. In Experiment 1, the electrode separation and the sequence duration were varied to clarify whether place cues help CI listeners to voluntarily segregate sounds and whether a two-stream percept needs time to build up. Results suggested that place cues can facilitate the segregation of sequential sounds if enough time is provided to build up a two-stream percept. In Experiment 2, the duration of the sequence was fixed, and only the electrode separation was varied to estimate the fission boundary. Most listeners were able to segregate the sounds for separations of three or more electrodes, and some listeners could segregate sounds coming from adjacent electrodes.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Paredes Gallardo, A. (Intern), Madsen, S. M. K. (Intern), Dau, T. (Intern), Marozeau, J. (Intern)
Number of pages: 13
Publication date: 2018
Main Research Area: Technical/natural sciences

Publication information
Journal: Trends in Hearing (Online)
Volume: 22
ISSN (Print): 2331-2165
Accuracy of averaged auditory brainstem response amplitude and latency estimates

Objective: The aims were to 1) establish which of the four algorithms for estimating residual noise level and signal-to-noise ratio (SNR) in auditory brainstem responses (ABRs) perform better in terms of post-average wave-V peak latency and amplitude errors and 2) determine whether SNR or noise floor is a better stop criterion where the outcome measure is peak latency or amplitude. Design: The performance of the algorithms was evaluated by numerical simulations using an ABR template combined with electroencephalographic (EEG) recordings obtained without sound stimulus. The suitability of a fixed SNR versus a fixed noise floor stop criterion was assessed when variations in the wave-V waveform shape reflecting inter-subject variation was introduced. Study sample: Over 100 hours of raw EEG noise was recorded from 17 adult subjects, under different conditions (e.g. sleep or movement). Results: ABR feature accuracy was similar for the four algorithms. However, it was shown that a fixed noise floor leads to higher ABR wave-V amplitude accuracy; conversely, a fixed SNR yields higher wave-V latency accuracy. Conclusion: Similar performance suggests the use of the less computationally complex algorithms. Different stop criteria are recommended if the ABR peak latency or the amplitude is the outcome measure of interest.
A Model of Electrically Stimulated Auditory Nerve Fiber Responses with Peripheral and Central Sites of Spike Generation

A computational model of cat auditory nerve fiber (ANF) responses to electrical stimulation is presented. The model assumes that (1) there exist at least two sites of spike generation along the ANF and (2) both an anodic (positive) and a cathodic (negative) charge in isolation can evoke a spike. A single ANF is modeled as a network of two exponential integrate-and-fire point-neuron models, referred to as peripheral and central axons of the ANF. The peripheral axon is excited by the cathodic charge, inhibited by the anodic charge, and exhibits longer spike latencies than the central axon; the central axon is excited by the anodic charge, inhibited by the cathodic charge, and exhibits shorter spike latencies than the peripheral axon. The model also includes subthreshold and suprathreshold adaptive feedback loops which continuously modify the membrane potential and can account for effects of facilitation, accommodation, refractoriness, and spike-rate adaptation in ANF. Although the model is parameterized using data for either single or paired pulse stimulation with monophasic rectangular pulses, it correctly predicts effects of various stimulus pulse shapes, stimulation pulse rates, and level on the neural response statistics. The model may serve as a framework to explore the effects of different stimulus parameters on psychophysical performance measured in cochlear implant listeners.
A Nonlinear Transmission Line Model of the Cochlea With Temporal Integration Accounts for Duration Effects in Threshold Fine Structure

For normal-hearing listeners, auditory pure-tone thresholds in quiet often show quasi periodic fluctuations when measured with a high frequency resolution, referred to as threshold fine structure. Threshold fine structure is dependent on the stimulus duration, with smaller fluctuations for short than for long signals. The present study demonstrates how this effect can be captured by a nonlinear and active model of the cochlear in combination with a temporal integration stage. Since this cochlear model also accounts for fine structure and connected level dependent effects, it is superior to filter-based approaches and hence allows the investigation of the contributions of cochlear- and retro-cochlear processing on behavioural data, including stimulus-duration dependent effects of threshold fine structure.
Assessing the efficacy of hearing-aid amplification using a phoneme test

Consonant-vowel (CV) perception experiments provide valuable insights into how humans process speech. Here, two CV identification experiments were conducted in a group of hearing-impaired (HI) listeners, using 14 consonants followed by the vowel /ɑ/. The CVs were presented in quiet and with added speech-shaped noise at signal-to-noise ratios of 0, 6, and 12 dB. The HI listeners were provided with two different amplification schemes for the CVs. In the first experiment, a frequency-independent amplification (flat-gain) was provided and the CVs were presented at the most-comfortable loudness level. In the second experiment, a frequency-dependent prescriptive gain was provided. The CV identification results showed that, while the average recognition error score obtained with the frequency-dependent amplification was lower than that obtained with the flat-gain, the main confusions made by the listeners on a token basis remained the same in a majority of the cases. An entropy measure and an angular distance measure were proposed to assess the highly individual effects of the frequency-dependent gain on the consonant confusions in the HI listeners. The results suggest that the proposed measures, in combination with a well-controlled phoneme speech test, may be used to assess the impact of hearing-aid signal processing on speech intelligibility.
Assessment of broadband SNR estimation for hearing aid applications

An accurate estimation of the broadband input signal-to-noise ratio (SNR) is a prerequisite for many hearing-aid algorithms. An extensive comparison of three SNR estimation algorithms was performed. Moreover, the influence of the duration of the analysis window on the SNR estimation performance was systematically investigated.

The most accurate approach utilized an estimation of the clean speech power spectral density (PSD) and the noisy speech power across a sliding window of 1280 ms and achieved an total SNR estimation error below 3 dB across a wide variety of background noises and input SNRs.

Auditory brainstem response latency in forward masking, a marker of sensory deficits in listeners with normal hearing thresholds

In rodent models, acoustic exposure too modest to elevate hearing thresholds can nonetheless cause auditory nerve fiber deafferentation, interfering with the coding of supra-threshold sound. Low-spontaneous rate nerve fibers, important for encoding acoustic information at supra-threshold levels and in noise, are more susceptible to degeneration than high-spontaneous rate fibers. The change in auditory brainstem response (ABR) wave-V latency with noise level has been shown to be associated with auditory nerve deafferentation. Here, we measured ABR in a forward masking paradigm and evaluated wave-V latency changes with increasing masker-to-probe intervals. In the same listeners, behavioral forward masking detection thresholds were measured. We hypothesized that 1) auditory nerve fiber deafferentation increases forward masking thresholds and increases wave-V latency and 2) a preferential loss of low-spontaneous rate fibers results in a faster recovery of wave-V latency as the slow contribution of these fibers is reduced. Results showed that in young audiometrically normal listeners, a larger change in wave-V latency with increasing masker-to-probe interval was positively correlated with the rate of change in forward masking detection thresholds. Although we cannot rule out central contributions, these findings are consistent with the hypothesis that auditory nerve fiber deafferentation occurs in humans and may predict how well individuals can hear in noisy environments. (C) 2017 Elsevier B.V. All rights reserved.
Cascaded Amplitude Modulations in Sound Texture Perception

Sound textures, such as crackling fire or chirping crickets, represent a broad class of sounds defined by their homogeneous temporal structure. It has been suggested that the perception of texture is mediated by time-averaged summary statistics measured from early auditory representations. In this study, we investigated the perception of sound
textures that contain rhythmic structure, specifically second-order amplitude modulations that arise from the interaction of different modulation rates, previously described as "beating" in the envelope-frequency domain. We developed an auditory texture model that utilizes a cascade of modulation filterbanks that capture the structure of simple rhythmic patterns. The model was examined in a series of psychophysical listening experiments using synthetic sound textures-stimuli generated using time-averaged statistics measured from real-world textures. In a texture identification task, our results indicated that second-order amplitude modulation sensitivity enhanced recognition. Next, we examined the contribution of the second-order modulation analysis in a preference task, where the proposed auditory texture model was preferred over a range of model deviants that lacked second-order modulation rate sensitivity. Lastly, the discriminability of textures that included second-order amplitude modulations appeared to be perceived using a time-averaging process. Overall, our results demonstrate that the inclusion of second-order modulation analysis generates improvements in the perceived quality of synthetic textures compared to the first-order modulation analysis considered in previous approaches.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: McWalter, R. I. (Intern), Dau, T. (Intern)
Number of pages: 12
Publication date: 2017
Main Research Area: Technical/natural sciences

Publication information
Journal: Frontiers in Neuroscience
Volume: 11
Article number: 485
ISSN (Print): 1662-4548
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 3.85 SJR 1.88 SNIP 1.087
Scopus rating (2015): SJR 2.022 SNIP 1.093 CiteScore 3.72
Scopus rating (2014): SJR 2.04 SNIP 1.097 CiteScore 3.84
Scopus rating (2013): SJR 2.068 SNIP 1.089 CiteScore 3.61
Scopus rating (2012): SJR 1.718 SNIP 1.004 CiteScore 3.25
Scopus rating (2011): SJR 1.707 SNIP 1.268
Scopus rating (2010): SJR 1.326 SNIP 0.709
Original language: English
Amplitude modulation, Auditory model, Auditory perception, Natural sound, Sound texture
Electronic versions:
fnins_11_00485.pdf
DOIs:
10.3389/fnins.2017.00485

Bibliographical note
© 2017 McWalter and Dau. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Source: FindIt
Source-ID: 2390293281
Publication: Research - peer-review › Journal article – Annual report year: 2017

Characterizing cochlear hearing impairment using advanced electrophysiological methods

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Encina Llamas, G. (Intern), Epp, B. (Intern), Dau, T. (Intern), Harte, J. (Intern)
Number of pages: 171
Publication date: 2017
Comparison of objective and subjective measures of cochlear compression in normal-hearing and hearing-impaired listeners

Among several behaviourial methods for estimating the basilar membrane input/output function, the temporal masking curve is the most popular. Distortion product otoacoustic emissions provide an objective measure for estimating cochlear compression. However, estimates from both methods have been poorly correlated in previous studies. We hypothesise that this could be due to the interplay between generator and reflection components in the recorded otoacoustic emissions. Here, compression estimates obtained with the two methods were compared at three audiometric frequencies (1, 2, and 4 kHz) for 10 normal-hearing and 6 hearing-impaired listeners. Distortion-product otoacoustic emissions were evoked using continuouslyswept tones, to separate the generator component and investigate the corresponding compressive characteristic. For hearing imaiired listeners, the estimates from the two methods were highly correlated.

Contribution of low- and high-frequency bands to binaural unmasking in hearing-impaired listeners

This study investigated the contribution of interaural timing differences (ITDs) in different frequency regions to binaural unmasking (BU) of speech. Speech reception thresholds (SRTs) and binaural intelligibility level differences (BILDs) were measured in two-talker babble in 6 young normal-hearing (NH) and 9 elderly hearing-impaired (HI) listeners with normal or close-to-normal hearing at and below 1.5 kHz. Target sentences were presented diotically, embedded in a stream of diotic or dichotic maskers. Both target and masker sentences were split into frequency regions above and below 1.25 kHz. In the dichotic listening conditions, the maskers were lateralized to the left side by introducing 0.68-ms ITDs in either the low-frequency band, the high-frequency band, or both bands simultaneously. BILDs were found to be similar in both listener groups when the ITDs were imposed on the low-frequency band only. ITDs in the high-frequency band alone did not produce any BILD in any of the groups. However, when ITDs were imposed in both frequency bands, the NH listeners yielded significantly greater BILDs than the HI listeners. The results suggest that, on a group level, HI listeners relied solely on ITDs in the low-frequency band while NH listeners were able to utilize envelope ITDs above 1.25 kHz to facilitate the BU of speech.
Data-driven approach for auditory profiling

Nowadays, the pure-tone audiogram is the main tool used to characterize hearing loss and to fit hearing aids. However, the perceptual consequences of hearing loss are typically not only associated with a loss of sensitivity, but also with a clarity loss that is not captured by the audiogram. A detailed characterization of hearing loss has to be simplified to efficiently explore the specific compensation needs of the individual listener. We hypothesized that any listener’s hearing can be characterized along two dimensions of distortion: type I and type II. While type I can be linked to factors affecting audibility, type II reflects non-audibility-related distortions. To test our hypothesis, the individual performance data from two previous studies were re-analyzed using an archetypal analysis. Unsupervised learning was used to identify extreme patterns in the data which form the basis for different auditory profiles. Next, a decision tree was determined to classify the listeners into one of the profiles. The new analysis provides evidence for the existence of four profiles in the data. The most significant predictors for profile identification were related to binaural processing, auditory non-linearity and speech-in-noise perception. The current approach is promising for analyzing other existing data sets in order to select the most relevant tests for auditory profiling.

Effect of musical training on pitch discrimination performance in older normal-hearing and hearing-impaired listeners

Hearing-impaired (HI) listeners, as well as elderly listeners, typically have a reduced ability to discriminate the fundamental frequency (F0) of complex tones compared to young normal-hearing (NH) listeners. Several studies have shown that musical training, on the other hand, leads to improved F0-discrimination performance for NH listeners. It is unclear whether a comparable effect of musical training occurs for listeners whose sensory encoding of F0 is degraded. To address this question, F0 discrimination was investigated for three groups of listeners (14 young NH, 9 older NH and 10 HI listeners), each including musicians and non-musicians, using complex tones that differed in harmonic content. Musical training significantly improved F0 discrimination for all groups of listeners, especially for complex tones containing low-numbered harmonics. In a second experiment, the sensitivity to temporal fine structure cues (TFS) was estimated in the same listeners. Although TFS cues were degraded for the two older groups of listeners, musicians showed better performance than non-musicians. Additionally, a significant correlation was obtained between F0-discrimination performance and sensitivity to TFS cues for complex tones with low and intermediate harmonic numbers. These findings suggest that musical training may enhance both sensory encoding of TFS cues and F0 discrimination in young and older listeners with or without hearing loss.
Effects of hearing-aid dynamic range compression on spatial perception in a reverberant environment

This study investigated the effects of fast-acting hearing-aid compression on normal-hearing and hearing-impaired listeners' spatial perception in a reverberant environment. Three compression schemes—independent compression at each ear, linked compression between the two ears, and "spatially ideal" compression operating solely on the dry source signal—were considered using virtualized speech and noise bursts. Listeners indicated the location and extent of their perceived sound images on the horizontal plane. Linear processing was considered as the reference condition. The results showed that both independent and linked compression resulted in more diffuse and broader sound images as well as internalization and image splits, whereby more image splits were reported for the noise bursts than for speech. Only the spatially ideal compression provided the listeners with a spatial percept similar to that obtained with linear processing. The same general pattern was observed for both listener groups. An analysis of the interaural coherence and direct-to-reverberant ratio suggested that the spatial distortions associated with independent and linked compression resulted from enhanced reverberant energy. Thus, modifications of the relation between the direct and the reverberant sound should be avoided in amplification strategies that attempt to preserve the natural sound scene while restoring loudness cues.
Effects of slow- and fast-acting compression on hearing impaired listeners' consonant-vowel identification in interrupted noise

There is conflicting evidence about the relative benefit of slow- and fast-acting compression for speech intelligibility. It has been hypothesized that fast-acting compression improves audibility at low signal-to-noise ratios (SNRs) but may distort the
speech envelope at higher SNRs. The present study investigated the effects of compression with nearly instantaneous attack time but either fast (10 ms) or slow (500 ms) release times on consonant identification in hearing-impaired listeners. Consonant-vowel speech tokens were presented at several presentation levels in two conditions: in the presence of interrupted noise and in quiet (with the compressor “shadow controlled” by the corresponding mixture of speech and noise). These conditions were chosen to disentangle the effects of consonant audibility and noise-induced forward masking on speech intelligibility. A small but systematic intelligibility benefit of fast-acting compression was found in both the quiet and the noisy conditions for the lower speech levels. No negative effects of fast-acting compression were observed when the speech level exceeded the level of the noise. These findings suggest that fast-acting compression provides an audibility benefit in fluctuating interferers as compared to slow-acting compression, while not substantially affecting the perception of consonants at higher SNRs.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Sonova U.S. Corporate Services
Authors: Kowalewski, B. (Intern), Zaar, J. (Intern), Fereczkowski, M. (Intern), MacDonald, E. (Intern), Strelcyk, O. (Ekstern), May, T. (Intern), Dau, T. (Intern)
Number of pages: 8
Publication date: 2017

Host publication information
Publisher: The Danavox Jubilee Foundation
Editors: Santurette, S., Dau, T., C.-Dalsgaard, J., Tranebjærg, L., Andersen, T., Poulsen, T.
ISBN (Electronic): 978-87-990013-6-1
Main Research Area: Technical/natural sciences
Conference: International Symposium on Auditory and Audiological Research, Nyborg, Denmark, 23/08/2017 - 23/08/2017
Electronic versions:
326_Article_PDF_file_1232_1_18_20171214.pdf
Publication: Research - peer-review › Article in proceedings – Annual report year: 2017

Evaluating a Loudspeaker-Based Virtual Sound Environment using Speech-on-Speech Masking

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Ahrens, A. (Intern), Marschall, M. (Intern), Dau, T. (Intern)
Pages: 1138-1141
Publication date: 2017

Host publication information
Title of host publication: Proceedings of DAGA 2017
Publisher: Deutsche Gesellschaft für Akustik e.V.
Main Research Area: Technical/natural sciences
Conference: 43. Jahrestagung für Akustik, Kiel, Germany, 06/03/2017 - 06/03/2017
Source: PublicationPreSubmission
Source-ID: 131769393
Publication: Research - peer-review › Article in proceedings – Annual report year: 2017

Exploiting Deep Neural Networks and Head Movements for Robust Binaural Localization of Multiple Sources in Reverberant Environments

This paper presents a novel machine-hearing system that exploits deep neural networks (DNNs) and head movements for robust binaural localization of multiple sources in reverberant environments. DNNs are used to learn the relationship between the source azimuth and binaural cues, consisting of the complete cross-correlation function (CCF) and interaural level differences (ILDs). In contrast to many previous binaural hearing systems, the proposed approach is not restricted to localization of sound sources in the frontal hemifield. Due to the similarity of binaural cues in the frontal and rear hemifields, front–back confusions often occur. To address this, a head movement strategy is incorporated in the localization model to help reduce the front–back errors. The proposed DNN system is compared to a Gaussian-mixture-model-based system that employs interaural time differences (ITDs) and ILDs as localization features. Our experiments show that the DNN is able to exploit information in the CCF that is not available in the ITD cue, which together with head movements substantially improves localization accuracies under challenging acoustic scenarios, in which multiple talkers and room reverberation are present.

General information
State: Published
Impact of Noise and Noise Reduction on Processing Effort: A Pupillometry Study

Speech perception in adverse listening situations can be exhausting. Hearing loss particularly affects processing demands, as it requires increased effort for successful speech perception in background noise. Signal processing in
hearing aids and noise reduction (NR) schemes aim to counteract the effect of noise and reduce the effort required for speech recognition in adverse listening situations. The present study examined the benefit of NR schemes, applying a combination of a digital NR and directional microphones, for reducing the processing effort during speech recognition. The effect of noise (intelligibility level) and different NR schemes on effort were evaluated by measuring the pupil dilation of listeners. In 2 different experiments, performance accuracy and peak pupil dilation (PPD) were measured in 24 listeners with hearing impairment while they performed a speech recognition task. The listeners were tested at 2 different signal to noise ratios corresponding to either the individual 50% correct (L50) or the 95% correct (L95) performance level in a 4-talker babble condition with and without the use of a NR scheme. In experiment 1, the PPD differed in response to both changes in the speech intelligibility level (L50 versus L95) and NR scheme. The PPD increased with decreasing intelligibility, indicating higher processing effort under the L50 condition compared with the L95 condition. Moreover, the PPD decreased when the NR scheme was applied, suggesting that the processing effort was reduced. In experiment 2, 2 hearing aids using different NR schemes (fast-acting and slow-acting) were compared. Processing effort changed as indicated by the PPD depending on the hearing aids and therefore on the NR scheme. Larger PPDs were measured for the slow-acting NR scheme. The benefit of applying an NR scheme was demonstrated for both L50 and L95, that is, a situation at which the performance level was at a ceiling. This opens the opportunity for new means of evaluating hearing aids in situations in which traditional speech reception measures are shown not to be sensitive. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Eriksholm Research Centre, Linköping University
Authors: Wendt, D. (Intern), Hietkamp, R. K. (Ekstern), Lunner, T. (Ekstern)
Pages: 690-700
Publication date: 2017
Main Research Area: Technical/natural sciences

Publication information
Journal: Ear and Hearing
Volume: 38
Issue number: 6
ISSN (Print): 0196-0202
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 2.97 SJR 1.865 SNIP 1.571
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 1.753 SNIP 2.008 CiteScore 2.94
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 1.93 SNIP 1.726 CiteScore 2.86
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 1.893 SNIP 2.109 CiteScore 3.18
ISI indexed (2013): ISI indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 1.918 SNIP 1.708 CiteScore 2.95
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 1.948 SNIP 1.736 CiteScore 2.85
ISI indexed (2011): ISI indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 2.059 SNIP 1.872
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 2.28 SNIP 1.925
BFI (2008): BFI-level 2
Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation

Previous research has reported effects of masker type and signal-to-noise ratio (SNR) on listening effort, as indicated by the peak pupil dilation (PPD) relative to baseline during speech recognition. At about 50% correct sentence recognition performance, increasing SNRs generally results in declining PPDs, indicating reduced effort. However, the decline in PPD over SNRs has been observed to be less pronounced for hearing-impaired (HI) compared to normal-hearing (NH) listeners. The presence of a competing talker during speech recognition generally resulted in larger PPDs as compared to the presence of a fluctuating or stationary background noise. The aim of the present study was to examine the interplay between hearing-status, a broad range of SNRs corresponding to sentence recognition performance varying from 0 to 100% correct, and different masker types (stationary noise and single-talker masker) on the PPD during speech perception. Twenty-five HI and 32 age-matched NH participants listened to sentences across a broad range of SNRs, masked with speech from a single talker (-25 dB to +15 dB SNR) or with stationary noise (-12 dB to +16 dB). Correct sentence recognition scores and pupil responses were recorded during stimulus presentation. With a stationary masker, NH listeners show maximum PPD across a relatively narrow range of low SNRs, while HI listeners show relatively large PPD across a wide range of ecological SNRs. With the single-talker masker, maximum PPD was observed in the mid-range of SNRs around 50% correct sentence recognition performance, while smaller PPDs were observed at lower and higher SNRs. Mixed-model ANOVAs revealed significant interactions between hearing-status and SNR on the PPD for both masker types. Our data show a different pattern of PPDs across SNRs between groups, which indicates that listening and the allocation of effort during listening in daily life environments may be different for NH and HI listeners.

General information

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, VU University Medical Centre, Linköping University, University of Nottingham, Vrije Universiteit Amsterdam
Authors: Ohlenforst, B. (Ekstern), Zekveld, A. A. (Ekstern), Lunner, T. (Ekstern), Wendt, D. (Intern), Naylor, G. (Ekstern), Wang, Y. (Ekstern), Versfeld, N. J. (Ekstern), Kramer, S. E. (Ekstern)
Pages: 68-79
Publication date: 2017
Main Research Area: Technical/natural sciences

Publication information

Journal: Hearing Research
Volume: 351
ISSN (Print): 0378-5955
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 3.12 SJR 1.512 SNIP 1.358
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 1.857 SNIP 1.478 CiteScore 3.28
Influence of a remote microphone on localization with hearing aids

When used with hearing aids (HA), the addition of a remote microphone (RM) may alter the spatial perception of the listener. First, the RM signal is presented diotically from the HAs. Second, the processing in the HA often delays the RM signal relative to the HA microphone signals. Finally, the level of the RM signal is independent of the distance from the RM to HA. The present study investigated localization performance of 15 normal-hearing and 9 hearing-impaired listeners under conditions simulating the use of an RM with a behind the ear (BTE) HA. Minimum audible angle discrimination around an average angle of 45° was measured for three sets of relative gains and seven sets of relative delays for a total of 21 conditions. In addition, a condition with just the simulated BTE HA signals was tested. Overall, for both groups, minimum audible angle discrimination was best when the relative RM gain was small (~3 and ~6 dB) and the delay was approximately 10-20 ms. Under these conditions, localization performance approached the level obtained in the BTE HA only condition.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, GN Hearing Danmark A/S
Authors: Selby, J. G. (Ekstern), Weisser, A. (Ekstern), MacDonald, E. (Intern)
Number of pages: 7
Influence of binary mask estimation errors on robust speaker identification

Missing-data strategies have been developed to improve the noise-robustness of automatic speech recognition systems in adverse acoustic conditions. This is achieved by classifying time-frequency (T-F) units into reliable and unreliable components, as indicated by a so-called binary mask. Different approaches have been proposed to handle unreliable feature components, each with distinct advantages. The direct masking (DM) approach attenuates unreliable T-F units in the spectral domain, which allows the extraction of conventionally used mel-frequency cepstral coefficients (MFCCs). Instead of attenuating unreliable components in the feature extraction front-end, full marginalization (FM) discards unreliable feature components in the classification back-end. Finally, bounded marginalization (BM) can be used to combine the evidence from both reliable and unreliable feature components during classification. Since each of these approaches utilizes the knowledge about reliable and unreliable feature components in a different way, they will respond differently to estimation errors in the binary mask. The goal of this study was to identify the most effective strategy to exploit knowledge about reliable and unreliable feature components in the context of automatic speaker identification (SID). A systematic evaluation under ideal and non-ideal conditions demonstrated that the robustness to errors in the binary mask varied substantially across the different missing-data strategies. Moreover, full and bounded marginalization showed complementary performances in stationary and non-stationary background noises and were subsequently combined using a simple score fusion. This approach consistently outperformed individual SID systems in all considered experimental conditions.
A recent study suggested that the most important factor for obtaining high speech intelligibility in noise with cochlear implant recipients is to preserve the low-frequency amplitude modulations of speech across time and frequency by, for example, minimizing the amount of noise in speech gaps. In contrast, other studies have argued that the transients provide the most information. Thus, the present study investigates the relative impact of these two factors in the framework of noise reduction by systematically correcting noise-estimation errors within speech segments, speech gaps, and the transitions between them. Speech intelligibility in noise was measured using a cochlear implant simulation tested on normal-hearing listeners. The results suggest that minimizing noise in the speech gaps can substantially improve intelligibility, especially in modulated noise. However, significantly larger improvements were obtained when both the noise in the gaps was minimized and the speech transients were preserved. These results imply that the correct identification of the boundaries between speech segments and speech gaps is the most important factor in maintaining high intelligibility in cochlear implants. Knowing the boundaries will make it possible for algorithms to both minimize the noise in the gaps and enhance the low frequency amplitude modulations.
Investigating time-efficiency of forward masking paradigms for estimating basilar membrane input-output characteristics

It is well known that pure-tone audiometry does not sufficiently describe individual hearing loss (HL) and that additional measures beyond pure-tone sensitivity might improve the diagnostics of hearing deficits. Specifically, forward masking experiments to estimate basilar membrane (BM) input-output (I/O) function have been proposed. However, such measures are very time consuming. The present study investigated possible modifications of the temporal masking curve (TMC) paradigm to improve time and measurement efficiency. In experiment 1, estimates of knee point (KP) and compression ratio (CR) of individual BM I/Os were derived without considering the corresponding individual "off-frequency" TMC. While accurate estimation of KPs was possible, it is difficult to ensure that the tested dynamic range is sufficient. Therefore, in experiment 2, a TMC-based paradigm, referred to as the "gap method", was tested. In contrast to the standard TMC paradigm, the masker level was kept fixed and the "gap threshold" was obtained, such that the masker just masks a low-level (12 dB sensation level) signal. It is argued that this modification allows for better control of the tested stimulus level range, which appears to be the main drawback of the conventional TMC method. The results from the present study were consistent with the literature when estimating KP levels, but showed some limitations regarding the estimation of the CR values. Perspectives and limitations of both approaches are discussed.
Lateralized speech perception with small interaural time differences in normal-hearing and hearing-impaired listeners

Spatial release from masking (SRM) elicited by interaural timing differences (ITDs) only can be almost normal for listeners with symmetrical hearing loss. This study investigated whether elderly hearing-impaired (HI) listeners still achieve similar SRMs as young normal-hearing (NH) listeners, when SRMs are elicited by small ITDs. Speech reception thresholds (SRTs) and SRM due to ITDs were measured over headphones for 10 young NH and 10 older HI listeners, who had normal or close-to-normal hearing below 1.5 kHz. Diotic target sentences were presented in diotic or dichotic speech-shaped noise or two-talker babble maskers. In the dichotic conditions, maskers were lateralized by delaying the masker waveforms in the left headphone channel. Multiple magnitudes of masker ITDs were tested in both noise conditions. Although deficits were observed in speech perception abilities in speechshaped noise and two-talker babble in terms of SRTs, HI listeners could utilize ITDs to a similar degree as NH listeners to facilitate the binaural unmasking of speech. A slight difference was observed between the group means when target and maskers were separated from each other by large ITDs, but not when separated by small ITDs. Thus, HI listeners do not appear to require larger ITDs than NH listeners do in order to receive a benefit from binaural unmasking.
Linear combination of auditory steady-state responses evoked by co-modulated tones

Up to medium intensities and in the 80–100-Hz region, the auditory steady-state response (ASSR) to a multi-tone carrier is commonly considered to be a linear sum of the dipoles from each tone specific ASSR generator. Here, this hypothesis was investigated when a unique modulation frequency is used for all carrier components. Listeners were presented with a co-modulated dual-frequency carrier (1 and 4 kHz), from which the modulator starting phase $U_i$ of the 1-kHz component was systematically varied. The results support the hypothesis of a linear superposition of the dipoles originating from different frequency specific ASSR generators.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Guérit, F. (Intern), Marozeau, J. (Intern), Epp, B. (Intern)
Pages: EL395-EL400
Publication date: 2017
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of the Acoustical Society of America
Volume: 142
Issue number: 4
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.63 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 0.848 SNIP 1.633
Web of Science (2008): Indexed yes
Modeling Speech Level as a Function of Background Noise Level and Talker-to-Listener Distance for Talkers Wearing Hearing Protection Devices

Purpose: Studying the variations in speech levels with changing background noise level and talker-to-listener distance for talkers wearing hearing protection devices (HPDs) can aid in understanding communication in background noise.

Method: Speech was recorded using an intra-aural HPD from 12 different talkers at 5 different distances in 3 different noise conditions and 2 quiet conditions.

Results: This article proposes models that can predict the difference in speech level as a function of background noise level and talker-to-listener distance for occluded talkers. The proposed model complements the existing model presented by Pelegrín-García, Smits, Brunskog, and Jeong (2011) and expands on it by taking into account the effects of occlusion and background noise level on changes in speech sound level.

Conclusions: Three models of the relationship between vocal effort, background noise level, and talker-to-listener distance for talkers wearing HPDs are presented. The model with the best prediction intervals is a talker-dependent model that requires the users' unoccluded speech level at 10 m as a reference. A model describing the relationship between speech level, talker-to-listener distance, and background noise level for occluded talkers could eventually be incorporated with radio protocols to transmit verbal communication only to an intended set of listeners within a given spatial range—this range being dependent on the changes in speech level and background noise level.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Ecole de Technologie Superieure, Universite de Montreal, Institut National de la Recherche Scientifique
Authors: Bouserhal, R. E. (Ekstern), Bockstael, A. (Ekstern), MacDonald, E. (Intern), Falk, T. H. (Ekstern), Voix, J. (Ekstern)
Pages: 3393-3403
Publication date: 2017
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of Speech, Language, and Hearing Research
Volume: 60
ISSN (Print): 1092-4388
Ratings:
BFI (2018): BFI-level 1
Musicians do not benefit from differences in fundamental frequency when listening to speech in competing speech backgrounds

Recent studies disagree on whether musicians have an advantage over non-musicians in understanding speech in noise. However, it has been suggested that musicians may be able to use differences in fundamental frequency (F0) to better understand target speech in the presence of interfering talkers. Here we studied a relatively large (N=60) cohort of young adults, equally divided between nonmusicians and highly trained musicians, to test whether the musicians were better able to understand speech either in noise or in a two-talker competing speech masker. The target speech and competing speech were presented with either their natural F0 contours or on a monotone F0, and the F0 difference between the target and masker was systematically varied. As expected, speech intelligibility improved with increasing F0 difference between the target and the two-talker masker for both natural and monotone speech. However, no significant intelligibility advantage was observed for musicians over non-musicians in any condition. Although F0 discrimination was significantly better for musicians than for non-musicians, it was not correlated with speech scores. Overall, the results do not support the hypothesis that musical training leads to improved speech intelligibility in complex speech or noise backgrounds.
Noise-robust cortical tracking of attended speech in real-world acoustic scenes

Selectively attending to one speaker in a multi-speaker scenario is thought to synchronize low-frequency cortical activity to the attended speech signal. In recent studies, reconstruction of speech from single-trial electroencephalogram (EEG) data has been used to decode which talker a listener is attending to in a two-talker situation. It is currently unclear how this generalizes to more complex sound environments. Behaviorally, speech perception is robust to the acoustic distortions that listeners typically encounter in everyday life, but it is unknown whether this is mirrored by a noise-robust neural tracking of attended speech. Here we used advanced acoustic simulations to recreate real-world acoustic scenes in the laboratory. In virtual acoustic realities with varying amounts of reverberation and number of interfering talkers, listeners selectively attended to the speech stream of a particular talker. Across the different listening environments, we found that the attended talker could be accurately decoded from single-trial EEG data irrespective of the different distortions in the acoustic input. For highly reverberant environments, speech envelopes reconstructed from neural responses to the distorted stimuli resembled the original clean signal more than the distorted input. With reverberant speech, we observed a late cortical response to the attended speech stream that encoded temporal modulations in the speech signal without its reverberant distortion. Single-trial attention decoding accuracies based on 40-50s long blocks of data from 64 scalp electrodes were equally high (80-90% correct) in all considered listening environments and remained statistically significant using down to 10 scalp electrodes and short
On-site and laboratory evaluations of soundscape quality in recreational urban spaces

Regulations for quiet urban areas are typically based on sound level limits alone. However, the nonacoustic context may be crucial for subjective soundscape quality. Aims: This study aimed at comparing the role of sound level and nonacoustic context for subjective urban soundscape assessment in the presence of the full on-site context, the visual context only, and without context. Materials and Methods: Soundscape quality was evaluated for three recreational urban spaces by using four subjective attributes: loudness, acceptance, stressfulness, and comfort. The sound level was measured at each site and simultaneous sound recordings were obtained. Participants answered questionnaires either on site or during laboratory listening tests, in which the sound recordings were presented with or without each site’s visual context consisting of two pictures. They rated the four subjective attributes along with their preference toward eight sound sources. Results: The sound level was found to be a good predictor of all subjective parameters in the laboratory, but not on site. Although all attributes were significantly correlated in the laboratory setting, they did not necessarily covary on site. Moreover, the availability of the visual context in the listening experiment had no significant effect on the ratings. The participants were overall more positive toward natural sound sources on site. Conclusion: The full immersion in the on-site nonacoustic context may be important when evaluating overall soundscape quality in urban recreational areas. Laboratory evaluations may not fully reflect how subjective loudness, acceptance, stressfulness, and comfort are affected by sound level.
Pitch matching in bimodal cochlear implant patients: Effects of frequency, spectral envelope, and level

This study systematically investigated the effects of frequency, level, and spectral envelope on pitch matching in twelve bimodal cochlear implant (CI) users. The participants were asked to vary the frequency and level of a pure or complex tone (adjustable sounds) presented in the nonimplanted ear to match the pitch and loudness of different reference stimuli presented to the implanted ear. Three reference sounds were used: single electrode pulse trains, pure tones, and piano notes. The data showed a significant effect of the frequency and complexity of the reference sounds. No significant effect of the level of the reference sounds was found. The magnitude of effect of frequency was compressed in the implanted ear: on average a difference of seven semitones in the non-implanted ear induced the same pitch change as a difference of 19 to 24 semitones for a stimulus presented to the implanted ear. The spectral envelope of the adjustable sound presented to the non-implanted ear also had a significant effect. The matched frequencies were higher by an average of six semitones for the pure tone compared to a complex tone. Overall, the CI listeners might have matched the stimuli based on timbre characteristics such as brightness.

General information
- State: Published
- Organisations: Department of Electrical Engineering, Hearing Systems, The Bionics Institute, The Bionic Ear Institute
- Authors: Maarefvand, M. (Ekstern), Blamey, P. J. (Ekstern), Marozeau, J. (Intern)
- Pages: 2854–2865
- Publication date: 2017
- Main Research Area: Technical/natural sciences

Publication information
- Journal: Journal of the Acoustical Society of America
- Volume: 142
- Issue number: 5
- ISSN (Print): 0001-4966
- Ratings:
  - BFI (2018): BFI-level 2
  - Web of Science (2018): Indexed yes
  - BFI (2017): BFI-level 2
  - Web of Science (2017): Indexed yes
  - BFI (2016): BFI-level 2
  - Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
  - Web of Science (2016): Indexed yes
  - BFI (2015): BFI-level 2
  - Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
  - Web of Science (2015): Indexed yes
  - BFI (2014): BFI-level 2
  - Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
  - Web of Science (2014): Indexed yes
  - BFI (2013): BFI-level 2
  - Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
  - ISI indexed (2013): ISI indexed yes
Predicting consonant recognition and confusions in normal-hearing listeners

The perception of consonants in background noise has been investigated in various studies and was shown to critically depend on fine details in the stimuli. In this study, a microscopic speech perception model is proposed that represents an extension of the auditory signal processing model by Dau, Kollmeier, and Kohlrausch [(1997). J. Acoust. Soc. Am. 102, 2892–2905]. The model was evaluated based on the extensive consonant perception data set provided by Zaar and Dau [(2015). J. Acoust. Soc. Am. 138, 1253–1267], which was obtained with normal-hearing listeners using 15 consonant-vowel combinations mixed with white noise. Accurate predictions of the consonant recognition scores were obtained across a large range of signal-to-noise ratios. Furthermore, the model yielded convincing predictions of the consonant confusion scores, such that the predicted errors were clustered in perceptually plausible confusion groups. The large predictive power of the proposed model suggests that adaptive processes in the auditory preprocessing in combination with a cross-correlation based template-matching back end can account for some of the processes underlying consonant perception in normal-hearing listeners. The proposed model may provide a valuable framework, e.g., for investigating the effects of hearing impairment and hearing-aid signal processing on phoneme recognition.
Predicting effects of additive noise and hearing-instrument signal processing on consonant recognition and confusions

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Zaar, J. (Intern), Dau, T. (Intern)
Number of pages: 1
Publication date: 2017
Event: Poster session presented at 9th Speech in Noise Workshop, Oldenburg, Germany.
Main Research Area: Technical/natural sciences
Electronic versions:
SpIN_poster_v02.pdf
Source: PublicationPreSubmission
Source-ID: 129907689
Publication: Research › Poster – Annual report year: 2017

This study investigated the influence of hearing-aid (HA) and cochlear-implant (CI) processing on consonant perception in normal-hearing (NH) listeners. Measured data were compared to predictions obtained with a speech perception model [Zaar and Dau (2017). J. Acoust. Soc. Am. 141, 1051–1064] that combines an auditory processing front end with a correlation-based template-matching back end. In terms of HA processing, effects of strong nonlinear frequency compression and impulse-noise suppression were measured in 10 NH listeners using consonant-vowel stimuli. Regarding CI processing, the consonant perception data from DiNino et al. [(2016). J. Acoust. Soc. Am. 140, 4404-4418] were considered, which were obtained with noise-vocoded vowel-consonant-vowel stimuli in 12 NH listeners. The inputs to the model were the same stimuli as were used in the corresponding experiments. The model predictions obtained for the two data sets showed a large agreement with the perceptual data both in terms of consonant recognition and confusions, demonstrating the model's sensitivity to supra-threshold effects of hearing-instrument signal processing on consonant perception. The results could be useful for the evaluation of hearing-instrument processing strategies, particularly when combined with simulations of individual hearing impairment.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Sonova AG, University of Washington
Authors: Zaar, J. (Intern), Schmitt, N. (Ekstern), Derleth, R. (Ekstern), DiNino, M. (Ekstern), Arenberg, J. G. (Ekstern), Dau, T. (Intern)
Number of pages: 11
Pages: 3216-26
Publication date: 2017
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of the Acoustical Society of America
Volume: 142
Issue number: 5
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Predicting effects of hearing-instrument signal processing on consonant recognition and confusions

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Zaar, J. (Intern), Dau, T. (Intern)
Pages: 1438-1441
Publication date: 2017

Host publication information
Title of host publication: Proceedings of DAGA 2017
Publisher: Deutsche Gesellschaft für Akustik e.V.
Main Research Area: Technical/natural sciences
Conference: 43. Jahrestagung für Akustik, Kiel, Germany, 06/03/2017 - 06/03/2017
Source: PublicationPreSubmission
Source-ID: 136921613
Publication: Research - peer-review › Article in proceedings – Annual report year: 2017

Prediction of speech intelligibility based on a correlation metric in the envelope power spectrum domain
A powerful tool to investigate speech perception is the use of speech intelligibility prediction models. Recently, a model was presented, termed correlation-based speech-based envelope power spectrum model (sEPSMcorr) [1], based on the auditory processing of the multi-resolution speech-based Envelope Power Spectrum Model (mr-sEPSM) [2], combined with the correlation back-end of the Short-Time Objective Intelligibility measure (STOI) [3]. The sEPSMcorr can accurately predict NH data for a broad range of listening conditions, e.g., additive noise, phase jitter and ideal binary mask processing.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Relano-Iborra, H. (Intern), May, T. (Intern), Zaar, J. (Intern), Scheidiger, C. (Intern), Dau, T. (Intern)
Number of pages: 1
Publication date: 2017
Event: Poster session presented at 40th MidWinter Meeting of the Association for Research in Otolaryngology, Baltimore, United States.
Main Research Area: Technical/natural sciences
Electronic versions: ARO2017_final_v2.pdf
Source: PublicationPreSubmission
Source-ID: 137071837
Publication: Research - peer-review › Poster – Annual report year: 2017

Preliminary investigation of the categorization of gaps and overlaps in turn-taking interactions: Effects of noise and hearing loss
Normal conversation requires interlocutors to monitor the ongoing acoustic signal to judge when it is appropriate to start talking. Categorical thresholds for gaps and overlaps in turn-taking interactions were measured for normal hearing and hearing-impaired listeners in both quiet and multitalker babble (+6 dB SNR). The slope of the categorization functions were significantly shallower for hearing impaired listeners and in the presence of background noise. Moreover, the categorization threshold for overlaps increased in background noise.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, GN ReSound A/S
Authors: Sørensen, A. J. (Intern), Weisser, A. (Ekstern), MacDonald, E. (Intern)
Number of pages: 8
Publication date: 2017

Host publication information
Title of host publication: Proceedings of the International Symposium on Auditory and Audiological Research
ISBN (Print): 978-87-990013-6-1
Preserving spatial perception in rooms using direct-sound driven dynamic range compression

Fast-acting hearing-aid compression systems typically distort the auditory cues involved in the spatial perception of sounds in rooms by enhancing low-level reverberant energy portions of the sound relative to the direct sound. The present study investigated the benefit of a direct-sound driven compression system that adaptively selects appropriate time constants to preserve the listener’s spatial impression. Specifically, fast-acting compression was maintained for time-frequency units dominated by the direct sound while the processing of the compressor was linearized for time-frequency units dominated by reverberation. This compression scheme was evaluated with normal-hearing listeners who indicated their perceived location and distribution of sound images in the horizontal plane for virtualized speech. The experimental results confirmed that both independent compression at each ear and linked compression across ears resulted in broader, sometimes internalized, sound images as well as image splits. In contrast, the linked direct-sound driven compression system provided the listeners with a spatial perception similar to that obtained with linear processing that served as the reference condition. The independent direct-sound driven compressor created a sense of movement of the sound between the two ears, suggesting that preserving the interaural level differences via linked compression is advantageous with the proposed direct-sound driven compression scheme.
Real-time estimation of eye gaze by in-ear electrodes

Cognitive control of a hearing aid is the topic for several ongoing studies. The relevance of these studies should be seen in the light of inadequate steering of current hearing aids. While most studies are concerned with auditory attention tracking from the electroencephalogram (EEG), a complimentary approach may be to use visual attention tracking to steer the devices. Visual attention may be characterized by gaze direction, which can be obtained by electrooculography (EOG). EOG may be recorded from electrodes placed in the ear canal, termed EarEOG. To test the comparison of conventional EOG and EarEOG recordings, we conducted two experiments with six subjects. In the first experiment, the subjects were instructed to follow a moving dot on the screen moving in large saccades. In the second experiment, there were five large targets, and within each target, the dot had minor movements. When comparing conventional EOG and EarEOG, correlations of 0.9 and 0.91 with standard deviations of 0.02 were obtained for the two experiments respectively. To assess the feasibility of using EarEOG in real-time, correlation between EarEOG and the timecourse of the dot position was performed. When both signals were filtered with the same real-time applicable filter, correlations of 0.83 and 0.85 with
standard deviations of 0.09 and 0.05 were found respectively to the two experiments. In conclusion, this study provides motivational aspects of using EarEOG to estimate eye gaze, as well as it identifies important future challenges in real-time applications to steer external devices such as a hearing aid.

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Eriksholm Research Centre
Authors: Favre-Félix, A. (Intern), Graversen, C. (Ekstern), Dau, T. (Intern), Lunner, T. (Ekstern)
Number of pages: 4
Pages: 4086-4089
Publication date: 2017

**Host publication information**

Title of host publication: Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society
Publisher: IEEE
Main Research Area: Technical/natural sciences
DOIs: 10.1109/EMBC.2017.8037754
Source: FindIt
Source-ID: 2389514027
Publication: Research - peer-review › Article in proceedings – Annual report year: 2017

**Relations Between Self-Reported Daily-Life Fatigue, Hearing Status, and Pupil Dilation During a Speech Perception in Noise Task**

People with hearing impairment are likely to experience higher levels of fatigue because of effortful listening in daily communication. This hearing-related fatigue might not only constrain their work performance but also result in withdrawal from major social roles. Therefore, it is important to understand the relationships between fatigue, listening effort, and hearing impairment by examining the evidence from both subjective and objective measurements. The aim of the present study was to investigate these relationships by assessing subjectively measured daily-life fatigue (self-report questionnaires) and objectively measured listening effort (pupillometry) in both normally hearing and hearing-impaired participants. Twenty-seven normally hearing and 19 age-matched participants with hearing impairment were included in this study. Two self-report fatigue questionnaires Need For Recovery and Checklist Individual Strength were given to the participants before the test session to evaluate the subjectively measured daily fatigue. Participants were asked to perform a speech reception threshold test with single-talker masker targeting a 50% correct response criterion. The pupil diameter was recorded during the speech processing, and we used peak pupil dilation (PPD) as the main outcome measure of the pupillometry. No correlation was found between subjectively measured fatigue and hearing acuity, nor was a group difference found between the normally hearing and the hearing-impaired participants on the fatigue scores. A significant negative correlation was found between self-reported fatigue and PPD. A similar correlation was also found between Speech Intelligibility Index required for 50% correct and PPD. Multiple regression analysis showed that factors representing "hearing acuity" and "self-reported fatigue" had equal and independent associations with the PPD during the speech in noise test. Less fatigue and better hearing acuity were associated with a larger pupil dilation. To the best of our knowledge, this is the first study to investigate the relationship between subjectively measured daily-life fatigue and an objective measure of pupil dilation, as an indicator of listening effort. These findings help to provide an empirical link between pupil responses, as observed in the laboratory, and daily-life fatigue. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Technical University of Denmark
Authors: Wang, Y. (Ekstern), Naylor, G. (Ekstern), Kramer, S. E. (Ekstern), Zekveld, A. A. (Ekstern), Wendt, D. (Intern), Ohlenforst, B. (Ekstern), Lunner, T. (Ekstern)
Pages: 573–582
Publication date: 2017
Main Research Area: Technical/natural sciences

**Publication information**

Journal: Ear and Hearing
Volume: 39
Issue number: 3
ISSN (Print): 0196-0202
Ratings:
Relationship between overall comfort and combined thermal and acoustic conditions in urban recreational spaces

General information
State: Published
Organisations: Department of Electrical Engineering, Acoustic Technology, Hearing Systems, Technical University of Denmark
Authors: Bjerre, L. C. (Ekstern), Santurette, S. (Intern), Jeong, C. (Intern)
Number of pages: 8
Publication date: 2017

Host publication information
Title of host publication: Proceedings of INTER-NOISE 2017
We appreciate the opportunity to respond to the comment on our study (Ohlenforst et al. 2016). The issue of concern to the letter writer, “The speech and the noise signals were separately compressed before the SNRs [signal-to-noise ratios] were computed based on the signal's root mean square values,” referred to a calibration process in which the compressor gain function, derived from the mixed signal, was applied separately to the speech signal and to the separate noise signals to quantify the output signal to noise ratio. This was done to investigate compressor functionality. However, to create the test stimuli presented to the participants, the compressor was applied to the mixed speech-plus-noise signal at a specified input signal to noise ratio (−4, −2, and 0 dB), not on the speech signal or the noise signals alone. That is, the actual stimulus processing was all done on the mixed speech-plus-noise signal, which we believe to be the method that best represents realistic hearing aid situations. We regret the confusion that this inadvertent omission has caused and would like to thank Dr. Leijon for bringing the issue to our attention.
Sound specificity effects in spoken word recognition: The effect of integrality between words and sounds

Recent evidence has shown that nonlinguistic sounds co-occurring with spoken words may be retained in memory and affect later retrieval of the words. This sound-specificity effect shares many characteristics with the classic voice-specificity effect. In this study, we argue that the sound-specificity effect is conditional upon the context in which the word and sound coexist. Specifically, we argue that, besides co-occurrence, integrality between words and sounds is a crucial factor in the emergence of the effect. In two recognition-memory experiments, we compared the emergence of voice and sound specificity effects. In Experiment 1, we examined two conditions where integrality is high. Namely, the classic voice-specificity effect (Exp. 1a) was compared with a condition in which the intensity envelope of a background sound was modulated along the intensity envelope of the accompanying spoken word (Exp. 1b). Results revealed a robust voice-specificity effect and, critically, a comparable sound-specificity effect: A change in the paired sound from exposure to test led to a decrease in word-recognition performance. In the second experiment, we sought to disentangle the contribution of integrality from a mere co-occurrence context effect by removing the intensity modulation. The absence of integrality led to the disappearance of the sound-specificity effect. Taken together, the results suggest that the assimilation of background sounds into memory cannot be reduced to a simple context effect. Rather, it is conditioned by the extent to which words and sounds are perceived as integral as opposed to distinct auditory objects.
Subcortical and cortical correlates of pitch discrimination: Evidence for two levels of neuroplasticity in musicians

Musicians are highly trained to discriminate fine pitch changes but the neural bases of this ability are poorly understood. It is unclear whether such training-dependent differences in pitch processing arise already in the subcortical auditory system or are linked to more central stages. To address this question, we combined psychoacoustic testing with functional MRI to measure cortical and subcortical responses in musicians and non-musicians during a pitch-discrimination task. First, we estimated behavioral pitch-discrimination thresholds for complex tones with harmonic components that were either resolved or unresolved in the auditory system. Musicians outperformed non-musicians, showing lower pitch-discrimination thresholds in both conditions. The same participants underwent task-related functional MRI, while they performed a similar pitch-discrimination task. To account for the between-group differences in pitch-discrimination, task difficulty was adjusted to each individual's pitch-discrimination ability. Relative to non-musicians, musicians showed increased neural responses to complex tones with either resolved or unresolved harmonics especially in right-hemispheric areas, comprising the right superior temporal gyrus, Heschl's gyrus, insular cortex, inferior frontal gyrus, and in the inferior colliculus. Both subcortical and cortical neural responses predicted the individual pitch-discrimination performance. However, functional activity in the inferior colliculus correlated with differences in pitch discrimination across all participants, but not within the musicians group alone. Only neural activity in the right auditory cortex scaled with the fine pitch-discrimination thresholds within the musicians. These findings suggest two levels of neuroplasticity in musicians, whereby training-dependent changes in pitch processing arise at the collicular level and are preserved and further enhanced in the right auditory cortex.
The role of temporal cues in voluntary stream segregation in cochlear implant listeners

Cochlear implant (CI) listeners experience difficulties in complex listening scenarios, where the auditory system is required to segregate a target signal from the competing sound sources. The present study investigated segregation abilities of CI listeners as a function of temporal cues and examined whether a two-stream percept occurs instantaneously or needs time to build up. CI users participated in a detection task where a sequence of regularly presented bursts of pulses (“B”) on a single electrode interleaved with an irregular sequence (“A”) presented on the same electrode with a different pulse rate. The pulse rate difference and the duration of the sequences were varied between trials. In half of the trials, a delay was added to the last burst of the regular A sequence and the listeners were asked to detect this delay. As the period between consecutive B bursts was jittered, time judgments between the A and B sequences did not provide a reliable cue to perform the task such that the segregation of A and B should improve performance. The results showed that performance improved with increasing rate differences and increasing sequence duration, suggesting that CI listeners can segregate sounds based on temporal cues and that this percept builds up over time.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Paredes Gallardo, A. (Intern), Madsen, S. M. K. (Intern), Dau, T. (Intern), Marozeau, J. (Intern)
Number of pages: 8
Publication date: 2017

Host publication information
Title of host publication: Proceedings of the International Symposium on Auditory and Audiological Research
ISBN (Print): 978-87-990013-6-1
Series: Proceedings of the International Symposium on Auditory and Audiological Research
Volume: 6
Main Research Area: Technical/natural sciences
Conference: International Symposium on Auditory and Audiological Research, Nyborg, Denmark, 23/08/2017 - 23/08/2017

Relations
Activities:
International Symposium on Auditory and Audiological Research
Source: PublicationPreSubmission
Source-ID: 143862188
Publication: Research - peer-review › Article in proceedings – Annual report year: 2018

The speech-based envelope power spectrum model (sEPSM) family: Development, achievements, and current challenges
Intelligibility models provide insights regarding the effects of target speech characteristics, transmission channels and/or auditory processing on the speech perception performance of listeners. In 2011, Jørgensen and Dau proposed the speech-based envelope power spectrum model [sEPSM, Jørgensen and Dau (2011). J. Acoust. Soc. Am. 130(3), 1475-1487]. It uses the signal-to-noise ratio in the modulation domain (SNRenv) as a decision metric and was shown to accurately predict the intelligibility of processed noisy speech. The sEPSM concept has since been applied in various subsequent models, which have extended the predictive power of the original model to a broad range of conditions. This contribution presents the most recent developments within the sEPSM "family": (i) A binaural extension, the B-sEPSM [Chabot-Leclerc et al. (2016). J. Acoust. Soc. Am. 140(1), 192-205] which combines better-ear and binaural unmasking processes and accounts for a large variety of spatial phenomena in speech perception; (ii) a correlation-based version [Relaño-Iborra et al. (2016). J. Acoust. Soc. Am. 140(4), 2670-2679] which extends the predictions of the early model to non-linear distortions, such as phase jitter and binary mask-processing; and (iii) a recent physiologically inspired extension, which allows to functionally account for effects of individual hearing impairment on speech perception.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Verbal attribute magnitude estimates of pulse trains across selectrode places and stimulation rates in cochlear implant listeners

For cochlear implant users, temporal and place cue are assumed to vary along two orthogonal perceptual dimensions linked to pitch height and timbre. Here, the effect of electrode place, pulse rate, and amplitude modulation frequency on those perceptual dimensions was investigated. Combinations of different electrode places with differing pulse rates or modulation frequencies were presented to the participants while they were asked to rate pitch height and sound quality using multiple verbal attributes. The results indicate that temporal and place cues induce two perceptual dimensions that can be both linked to pitch and timbre.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Lamping, W. (Intern), Santurette, S. (Intern), Marozeau, J. (Intern)
Pages: 215-222
Publication date: 2017

Methodology and apparatus for determining psychoacoustical threshold curves

The present invention relates in a first aspect to a method of determining a psychoacoustical threshold curve by selectively varying a first parameter and a second parameter of an auditory stimulus signal applied to a test subject/listener. The methodology comprises steps of determining a two-dimensional boundary region surrounding an a priori estimated placement of the psychoacoustical threshold curve to form a predetermined two-dimensional response space comprising a positive response region at a first side of the a priori estimated psychoacoustical threshold curve and a negative response region at a second and opposite side of the a priori estimated psychoacoustical threshold curve. A series of auditory stimulus signals in accordance with the respective parameter pairs are presented to the listener through a sound reproduction device and the listener's detection of a predetermined attribute feature of the auditory stimulus signals is recorded such that a stimuli path through the predetermined two-dimensional response space is traversed. The psychoacoustical threshold curve is computed based on at least a subset of the recorded parameter pairs.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Fereczkowski, M. (Intern), MacDonald, E. (Intern), Dau, T. (Intern)
Publication date: 31 Mar 2016

Publication information
A correlation metric in the envelope power spectrum domain for speech intelligibility prediction

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Iborra, H. R. (Intern), May, T. (Intern), Zaar, J. (Intern), Scheidiger, C. (Intern), Dau, T. (Intern)
Publication date: 2016
Event: Poster session presented at ARCHES/ICANHEAR 2016, Zurich, Switzerland.
Main Research Area: Technical/natural sciences
Electronic versions:
ARCHES_poster_final3.pdf

Relations
Activities:
A correlation metric in the envelope power spectrum domain for speech intelligibility prediction
Publication: Research - peer-review › Poster – Annual report year: 2017

A modular guitar for teaching musical acoustics
In order to keep students activated in a course on musical acoustics, they were asked to build a modular guitar, designed to be updated throughout the course. In the first stage, dedicated to the physics of strings, a guitar was made out of three strings attached to a long piece of wood. The students measured the effect of the place of plucking on the mode of the vibrations of the strings. The second stage was dedicated to the acoustic resonances. Using a laser cutter, the students built a wooden box that was coupled to their guitar using straps. New acoustical measurements were made to study the effect of the shape of the resonator on the spectrum of the sound. In the third stage, as the different tuning systems were learned, the students built a fingerboard with the appropriated positions of the frets. In the last stage, the students have implemented some digital effects and tested them on their guitar using a piezo-electrical pickup. As nothing was glued, the students were able to easily change each part of the guitar (resonator, sound hole, fret positions, microphone, ...) in order to experience their direct effect and their interactions.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Marozeau, J. (Intern)
Pages: 3196-3196
Publication date: 2016
Main Research Area: Technical/natural sciences
Publication information
Journal: Journal of the Acoustical Society of America
Volume: 140
Issue number: 4
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
A new procedure for automatic fitting of the basilar-membrane input-output function to individual behavioral data. The basilar membrane input-output function (BM I/O) in a healthy cochlea is highly nonlinear. One of the consequences of sensorineural hearing loss (SNHL) is a partial or full loss of this nonlinearity. Behavioral estimates of the individual BM I/O
can be useful for modeling the impaired auditory system and, potentially, for clinical diagnostics. Computational algorithms are available that mimic the functioning of the nonlinear cochlear processing. One such algorithm is the dual resonance non-linear (DRNL) filterbank [6]. Its parameters can be modified to account for individual hearing loss, e.g., based on behavioral, temporal masking curves (TMC) data. This approach was used within the framework of the computational auditory signal-processing and perception (CASP) model to account for various aspects of SNHL [4]. However, due to the computational complexity, on-line fitting of the DRNL parameters is difficult. Until recently, the parameters were manually adjusted and the fitting process was indirect. A new approach is described here, based on a search through a lookup table of pre-computed filterbank input-output functions. The aim of this approach is to provide a fast, stable, and more objective fitting procedure.
Assessing the contribution of binaural cues for apparent source width perception via a functional model

In echoic conditions, sound sources are not perceived as point sources but appear to be expanded. The expansion in the horizontal dimension is referred to as apparent source width (ASW). To elicit this perception, the auditory system has access to fluctuations of binaural cues, the interaural time differences (ITDs), interaural level differences (ILDs) and the interaural coherence (IC). To quantify their contribution to ASW, a functional model of ASW perception was exploited using the TWO!EARS auditory-front-end (AFE) toolbox. The model determines the leftmost and rightmost boundary of a sound source using a statistical representation of ITDs and ILDs based on percentiles integrated over time and frequency. The model's performance was evaluated against psychoacoustic data obtained with noise, speech and music signals in loudspeaker-based experiments. A robust model prediction of ASW was achieved using a cross-correlation based estimation with either IC or ITDs, in contrast to a combination of ITDs and ILDs where the performance slightly decreased.

Auditory features in consonant perception - a modeling perspective

Most state-of-the-art hearing aids apply multi-channel dynamic-range compression (DRC). Such designs have the potential to emulate, at least to some degree, the processing that takes place in the healthy auditory system. One way to assess hearing-aid performance is to measure speech intelligibility. However, due to the complexity of speech and its robustness to spectral and temporal alterations, the effects of DRC on speech perception have been mixed and controversial. The goal of the present study was to obtain a clearer understanding of the interplay between hearing loss...
and DRC by means of auditory modeling. Inspired by the work of Edwards (2002), we studied the effects of DRC on a set of relatively basic outcome measures, such as forward masking functions (Glasberg and Moore, 1987) and spectral masking patterns (Moore et al., 1998), obtained at several masker levels and frequencies. Outcomes were simulated using the auditory processing model of Jepsen et al. (2008) with the front end modified to include effects of hearing impairment and DRC. The results were compared to experimental data from normal-hearing and hearing-impaired listeners.

**General information**

State: Published  
Organisations: Department of Electrical Engineering, Hearing Systems, Sonova U.S. Corporate Services  
Authors: Kowalewski, B. (Intern), MacDonald, E. (Intern), Strelcyk, O. (Ekstern), Dau, T. (Intern)  
Number of pages: 8  
Publication date: 2016

**Host publication information**

Title of host publication: Proceedings of ISAAR 2015  
Editors: Santurette, S., Dau, T., Dalsgaard, J. C., Tranebjærg, L., Andersen, T.  
ISBN (Print): 978-87-990013-5-4  
Main Research Area: Technical/natural sciences  
Conference: 5th International Symposium on Auditory and Audiological Research, Nyborg, Denmark, 26/08/2015 - 26/08/2015  
Source: PublicationPreSubmission  
Source-ID: 123395878  
Publication: Research - peer-review › Article in proceedings – Annual report year: 2016

**Auditory profiling and hearing-aid satisfaction in hearing-aid candidates**

Hearing-impaired (HI) listeners often complain about difficulties communicating in the presence of background noise, although audibility may be restored by a hearing-aid (HA). The audiogram typically forms the basis for HA fitting, i.e. people with similar audiograms are given the same prescription by default. This study aimed at identifying clinically relevant tests that may serve as an informative addition to the audiogram and which may relate more directly to HA satisfaction than the audiogram does. METHODS: A total of 29 HI and 26 normal-hearing listeners performed tests of spectral and temporal resolution, binaural hearing, speech intelligibility in stationary and fluctuating noise and a working-memory test. Six weeks after HA fitting, the HI listeners answered a questionnaire evaluating HA treatment. RESULTS: No other measures than masking release between fluctuating and stationary noise correlated significantly with audibility. The HI listeners who obtained the least advantage from fluctuations in background noise in terms of speech intelligibility experienced greater HA satisfaction. CONCLUSION: HI listeners have difficulties in different hearing domains that are not predictable from their audiogram. Measures of temporal resolution or speech perception in both stationary and fluctuating noise could be relevant measures to consider in an extended auditory profile. FUNDING: The study was supported by Grosserer L.F. Foghts Fond. TRIAL REGISTRATION: The protocol was approved by the Science Ethics Committee of the Capital Region of Denmark (reference H-3-2013-004).

**General information**

State: Published  
Organisations: Department of Electrical Engineering, Hearing Systems, University of Copenhagen  
Authors: Thorup, N. (Ekstern), Santurette, S. (Intern), Jørgensen, S. (Intern), Kjærbøl, E. (Ekstern), Dau, T. (Intern), Friis, M. (Ekstern)  
Number of pages: 5  
Publication date: 2016  
Main Research Area: Technical/natural sciences

**Publication information**

Journal: Danish Medical Journal  
Volume: 63  
Issue number: 10  
Article number: A5275  
ISSN (Print): 2245-1919  
Ratings:  
BFI (2018): BFI-level 1  
Web of Science (2018): Indexed yes  
BFI (2017): BFI-level 1  
Web of Science (2017): Indexed Yes  
BFI (2016): BFI-level 1
Can place-specific cochlear dispersion be represented by auditory steady-state responses?

The present study investigated to what extent properties of local cochlear dispersion can be objectively assessed through auditory steady-state responses (ASSR). The hypothesis was that stimuli compensating for the phase response at a particular cochlear location generate a maximally modulated basilar membrane (BM) response at that BM position, due to the large "within-channel" synchrony of activity. This would lead, in turn, to a larger ASSR amplitude than other stimuli of corresponding intensity and bandwidth. Two stimulus types were chosen: 1] Harmonic tone complexes consisting of equal-amplitude tones with a starting phase following an algorithm developed by Schroeder [IEEE Trans. Inf. Theory 16, 85-89 (1970)] that have earlier been considered in behavioral studies to estimate human auditory filter phase responses; and 2] simulations of auditory-filter impulse responses (IR). In both cases, also the temporally reversed versions of the stimuli were considered. The ASSRs obtained with the Schroeder tone complexes were found to be dominated by "across-channel" synchrony and, thus, do not reflect local place-specific information. In the case of the more frequency-specific stimuli, no significant differences were found between the responses to the IR and its temporally reversed counterpart. Thus, whereas ASSRs to narrowband stimuli have been used as an objective indicator of frequency-specific hearing sensitivity, the method does not seem to be sensitive enough to reflect local cochlear dispersion.
Sensory Systems, Auditory steady-state response, Cochlear dispersion, Cochlear filter, Cochlear mechanics, Electrophysiology
Clustering of Cochlear Oscillations in Frequency Plateaus as a Tool to Investigate SOAE Generation

Spontaneous otoacoustic emissions (SOAE) reflect the net effect of self-sustained activity in the cochlea, but do not directly provide information about the underlying mechanism and place of origin within the cochlea. The present study investigates if frequency plateaus as found in a linear array of coupled oscillators (OAM) [7] are also found in a transmission line model (TLM) which is able to generate realistic SOAEs [2] and if these frequency plateaus can be used to explain the formation of SOAEs. The simulations showed a clustering of oscillators along the simulated basilar membrane. Both, the OAM and the TLM show traveling-wave behavior along the oscillators coupled into one frequency plateau. While in the TLM roughness is required in order to produce SOAEs, no roughness is required to trigger frequency plateaus in the linear array of oscillators. The formation of frequency plateaus as a consequence of coupling between neighboring active oscillators might be the mechanism underlying SOAEs.
Comparing eye tracking with electrooculography for measuring individual sentence comprehension duration

The aim of this study was to validate a procedure for performing the audio-visual paradigm introduced by Wendt et al. (2015) with reduced practical challenges. The original paradigm records eye fixations using an eye tracker and calculates the duration of sentence comprehension based on a bootstrap procedure. In order to reduce practical challenges, we first reduced the measurement time by evaluating a smaller measurement set with fewer trials. The results of 16 listeners showed effects comparable to those obtained when testing the original full measurement set on a different collective of listeners. Secondly, we introduced electrooculography as an alternative technique for recording eye movements. The correlation between the results of the two recording techniques (eye tracker and electrooculography) was $r = 0.97$, indicating that both methods are suitable for estimating the processing duration of individual participants. Similar changes in processing duration arising from sentence complexity were found using the eye tracker and the electrooculography procedure. Thirdly, the time course of eye fixations was estimated with an alternative procedure, growth curve analysis, which is more commonly used in recent studies analyzing eye tracking data. The results of the growth curve analysis were compared with the results of the bootstrap procedure. Both analysis methods show similar processing durations.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Oldenburg
Authors: Müller, J. A. (Ekstern), Wendt, D. (Intern), Kollmeier, B. (Ekstern), Brand, T. (Ekstern)
Publication date: 2016
Main Research Area: Technical/natural sciences

Publication information
Journal: P L o S One
Volume: 11
Issue number: 10
Article number: e016462
ISSN (Print): 1932-6203
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2016): CiteScore 3.11 SJR 1.201 SNIP 1.092
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 1.414 SNIP 1.131 CiteScore 3.32
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 1.545 SNIP 1.141 CiteScore 3.54
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 1.74 SNIP 1.147 CiteScore 3.94
ISI indexed (2013): ISI indexed yes
Comparing the influence of spectro-temporal integration in computational speech segregation

The goal of computational speech segregation systems is to automatically segregate a target speaker from interfering maskers. Typically, these systems include a feature extraction stage in the front-end and a classification stage in the back-end. A spectrotemporal integration strategy can be applied in either the frontend, using the so-called delta features, or in the back-end, using a second classifier that exploits the posterior probability of speech from the first classifier across a spectro-temporal window. This study systematically analyzes the influence of such stages on segregation performance, the error distributions and intelligibility predictions. Results indicated that it could be problematic to exploit context in the back-end, even though such a spectro-temporal integration stage improves the segregation performance. Also, the results emphasized the potential need of a single metric that comprehensively predicts computational segregation performance and correlates well with intelligibility. The outcome of this study could help to identify the most effective spectro-temporal integration strategy for computational segregation systems.

General information

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Bentsen, T. (Intern), May, T. (Intern), Kressner, A. A. (Intern), Dau, T. (Intern)
Number of pages: 5
Publication date: 2016

Host publication information

Title of host publication: Proceedings of Interspeech 2016
Publisher: International Speech Communication Association
Main Research Area: Technical/natural sciences
Complex-Tone Pitch Discrimination in Listeners With Sensorineural Hearing Loss

Physiological studies have shown that noise-induced sensorineural hearing loss (SNHL) enhances the amplitude of envelope coding in auditory-nerve fibers. As pitch coding of unresolved complex tones is assumed to rely on temporal envelope coding mechanisms, this study investigated pitch-discrimination performance in listeners with SNHL. Pitch-discrimination thresholds were obtained for 14 normal-hearing (NH) and 10 hearing-impaired (HI) listeners for sine-phase (SP) and random-phase (RP) complex tones. When all harmonics were unresolved, the HI listeners performed, on average, worse than NH listeners in the RP condition but similarly to NH listeners in the SP condition. The increase in pitch-discrimination performance for the SP relative to the RP condition (F0DL ratio) was significantly larger in the HI as compared with the NH listeners. Cochlear compression and auditory-filter bandwidths were estimated in the same listeners. The estimated reduction of cochlear compression was significantly correlated with the increase in the F0DL ratio, while no correlation was found with filter bandwidth. The effects of degraded frequency selectivity and loss of compression were considered in a simplified peripheral model as potential factors in envelope enhancement. The model revealed that reducing cochlear compression significantly enhanced the envelope of an unresolved SP complex tone, while not affecting the envelope of a RP complex tone. This envelope enhancement in the SP condition was significantly correlated with the increased pitch-discrimination performance for the SP relative to the RP condition in the HI listeners.
Complex-tone pitch representations in the human auditory system.
Understanding how the human auditory system processes the physical properties of an acoustical stimulus to give rise to a pitch percept is a fascinating aspect of hearing research. Since most natural sounds are harmonic complex tones, this work focused on the nature of pitch-relevant cues that are necessary for the auditory system to retrieve the pitch of complex sounds. The existence of different pitch-coding mechanisms for low-numbered (spectrally resolved) and high-numbered (unresolved) harmonics was investigated by comparing pitch-discrimination performance across different cohorts of listeners, specifically those showing enhanced pitch cues (i.e., musicians) and those typically having disrupted pitch cues (i.e., hearing-impaired listeners). In particular, two main topics were addressed: the relative importance of resolved and unresolved harmonics for normal-hearing (NH) and hearing-impaired (HI) listeners and the effect of musical training for pitch discrimination of complex tones with resolved and unresolved harmonics. Concerning the first topic, behavioral and modeling results in listeners with sensorineural hearing loss (SNHL) indicated that temporal envelope cues of complex tones with unresolved harmonics may be enhanced relative to NH listeners at the output of peripheral auditory filters. This enhancement of temporal envelope coding was found to be ascribed to a reduction of cochlear compression. Since frequency selectivity and temporal fine structure (TFS) cues are known to be degraded in listeners with SNHL, it is likely that HI listeners rely on the enhanced envelope cues to retrieve the pitch of unresolved harmonics. Hence, the relative importance of pitch cues may be altered in HI listeners, whereby envelope cues may be used instead of TFS cues to obtain a similar performance in pitch discrimination to that of NH listeners. In the second part of this work, behavioral and objective measures of pitch discrimination were carried out in musicians and non-musicians. Musicians showed an increased pitch-discrimination performance relative to non-musicians for both resolved and unresolved harmonics, although their benefit was larger for the resolved harmonics. Additionally, task-evoked pupil responses were recorded as an indicator of processing effort while listeners performed a pitch-discrimination task. Although the difficulty of the task was adjusted for each participant to compensate for the individual pitch-discrimination abilities, the musically trained listeners still allocated lower processing effort than did the non-musicians to perform the task at the same performance level. This finding suggests an enhanced pitch representation along the auditory system in musicians, possibly as a result of training, which seemed to be specific to the stimuli containing resolved harmonics.

Finally, a functional magnetic resonance imaging paradigm was used to examine the response of the auditory cortex to resolved and unresolved harmonics in musicians and non-musicians. The neural responses in musicians were enhanced relative to the non-musicians for both resolved and unresolved harmonics in the right auditory cortex, right frontal regions and inferior colliculus. However, the increase in neural activation in the right auditory cortex of musicians was predictive of the increased pitch-discrimination performance only for resolved harmonics. These results suggest a training-dependent effect in musicians that is partially specific to the resolved harmonics.
Cross-correlation model of interaural time difference coding in listeners with bilateral cochlear implants

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Joshi, S. N. (Intern), Dau, T. (Intern), Epp, B. (Intern)
Number of pages: 1
Publication date: 2016
Event: Poster session presented at 5th Joint Meeting of the Acoustical Society of America and Acoustical Society of Japan, Honolulu, Hawaii, United States.
Main Research Area: Technical/natural sciences
Electronic versions:
Cross_correlation_model.pdf

Bibliographical note
The original title of the submitted and indexed abstract was "Modelling the effect of pulse-rate on coding of interaural time differences in listeners with cochlear implants" (http://asa.scitation.org/doi/abs/10.1121/1.4970392) but was changed to reflect the further work that was completed before the conference.
Source: PublicationPreSubmission
Source-ID: 128333961
Publication: Research › Poster – Annual report year: 2017

Effect of modulation depth, frequency, and intermittence on wind turbine noise annoyance
Amplitude modulation (AM) may be an important factor for the perceived annoyance of wind turbine noise (WTN). Two AM types, typically referred to as “normal AM” (NAM) and “other AM” (OAM), characterize WTN AM, OAM corresponding to having intermittent periods with larger AM depth in lower frequency regions than NAM. The extent to which AM depth, frequency, and type affect WTN annoyance remains uncertain. Moreover, the temporal variations of WTN AM have often not been considered. Here, realistic stimuli accounting for such temporal variations were synthesized such that AM depth, frequency, and type, while determined from real on-site recordings, could be varied systematically. Listening tests with both original and synthesized stimuli showed that a reduction in mean AM depth across the spectrum led to a significant decrease in annoyance. When the spectrotemporal characteristics of the original far-field stimuli and the temporal AM variations were taken into account, the effect of AM frequency remained limited and the presence of intermittent OAM periods did not affect annoyance. These findings suggest that, at a given overall level, the AM depth of NAM periods is the most crucial AM parameter for WTN annoyance.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Acoustic Technology
Authors: Ioannidou, C. (Intern), Santurette, S. (Intern), Jeong, C. (Intern)
Pages: 1241–1251
Publication date: 2016
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of the Acoustical Society of America
Volume: 139
Issue number: 3
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Effects of dynamic-range compression on temporal acuity

Some of the challenges that hearing-aid listeners experience with speech perception in complex acoustic environments may originate from limitations in the temporal processing of sounds. To systematically investigate the influence of hearing impairment and hearing-aid signal processing on temporal processing, temporal modulation transfer functions (TMTFs) and "supra-threshold" modulation-depth discrimination (MDD) thresholds were obtained in normal-hearing (NH) and hearing-impaired (HI) listeners with and without wide-dynamic range compression (WDRC). The TMTFs were obtained using tonal carriers of 1 and 5 kHz and modulation frequencies from 8 to 256 Hz. MDD thresholds were obtained using a reference modulation depth of -15 dB. A compression ratio of 2:1 was chosen. The attack and release time constants were 10 and 60 ms, respectively. For both carrier frequencies the TMTF thresholds decreased with the physical compression of the modulation depth due to the WDRC. Indications of reduced temporal resolution in the HI listeners were observed in the TMTF patterns for the 5 kHz carrier. Significantly higher MDD thresholds were found for the HI group relative to the NH group. No relationship was found between the MDD thresholds and the TMTF threshold. These findings indicate that the two measures may represent different aspects of temporal processing.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Widex A/S
Authors: Wiinberg, A. (Intern), Jepsen, M. L. (Ekstern), Epp, B. (Intern), Dau, T. (Intern)
Number of pages: 8
Publication date: 2016

Host publication information
Title of host publication: Proceedings of ISAAR 2015
Editors: Santurette, S., Dau, T., Dalsgaard, J. C., Tranebjærg, T., Andersen, T.
Main Research Area: Technical/natural sciences
Conference: 5th International Symposium on Auditory and Audiological Research, Nyborg, Denmark, 26/08/2015 - 26/08/2015
Source: PublicationPreSubmission
Source-ID: 123359348
Publication: Research - peer-review › Article in proceedings – Annual report year: 2016

Effets d'une compression cochléaire et d'une sélectivité en fréquence réduites sur la discrimination de la hauteur de tons complexes

Des études physiologiques ont montré qu'une perte auditive neurosensorielle (PANS) augmente l'amplitude de l'enveloppe temporelle du signal (ci-après, enveloppe) dans les fibres nerveuses auditives. La perception de la hauteur des tons complexes aux harmoniques non résolus reposant sur des mécanismes de codage de l'enveloppe, cette étude examine si des changements dans le traitement de l'enveloppe chez des sujets avec PANS ont des conséquences sur la discrimination de la hauteur. Des seuils de discrimination de fréquence fondamentale (SDF0) sont tout d'abord obtenus chez 14 sujets à audition normale et 10 sujets avec PANS pour des tons complexes dont les harmoniques sont ajoutées soit en phase sinusoïdale (PS) soit en phase aléatoire (PA). Pour des harmoniques non résolus, une PANS entraîne des SDF0 plus élevés que chez les personnes à audition normale dans la condition PA, mais similaires dans la condition PS. La compression cochléaire et la bande passante des filtres auditifs sont ensuite estimés chez les mêmes sujets. Les résultats démontrent une corrélation significative entre la réduction de la compression cochléaire et le rapport entre les SDF0 pour les conditions PA et PS. Les effets d'une dégradation de sélectivité en fréquence et d'une perte de compression sont enfin pris en compte comme facteurs potentiels de l'augmentation de l'enveloppe dans un modèle simplifié de la périphérie auditive. Ces simulations suggèrent que la réduction de la compression cochléaire et l'augmentation de la largeur des filtres auditifs améliorent sensiblement la représentation de l'enveloppe pour des harmoniques en PS, tout en l'affectant à peine pour des harmoniques en PA. Dans l'ensemble, les résultats comportementaux et de modélisation indiquent que la réduction de compression cochléaire est le facteur dominant dans l'augmentation de l'amplitude de l'enveloppe des tons complexes non résolu en PS, conduisant à une discrimination accrue de leur hauteur chez les personnes avec PANS

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Santurette, S. (Intern), Bianchi, F. (Intern), Fereczkowski, M. (Intern), Zaar, J. (Intern), Dau, T. (Intern)
Pages: 805-810
Publication date: 2016

Host publication information
Title of host publication: Proceedings of the French Acoustics Congress
Ellipsoidal reflector for measuring otoacoustic emissions

Otoacoustic emissions (OAEs) are low-intensity sounds present in the ear canal, generated by mechanical processing in the cochlear in the inner ear. OAEs provide a noninvasive technique to sense the mechanical processing of sound in the inner ear. These signals are commonly measured by placing a miniature microphone into the ear canal of the listener. Such small microphones have however a self-noise at frequencies below 1 kHz that is comparable in intensity with typical intensities of OAEs at frequencies above 1 kHz. Due to this fact, not much is known about the presence or absence of OAEs, and especially SOAE at these low frequencies. In addition, blocking of the ear canal changes the impedance of the middle ear, potentially changing the transmission of acoustical energy from the inner ear to the ear canal, hampering the interpretation of the data in terms of normal listening conditions with open ear canal. This study presents the design and evaluation of a truncated prolate ellipsoidal reflector in combination with a large-diaphragm low-noise microphone to measure OAEs in the open ear canal of human listeners. The reflector was designed to gain information about BM processing at low frequencies where miniature microphones are not easily applicable. Acoustical evaluation of the reflector shows a focusing effect of sound from one focal point into the other focal point. Partial removal of elements of the reflector allow to control multiple reflections between the microphone membrane and the ear canal. Spontaneous-and distortion-product OAEs show similar amplitudes and an improved noise floor at frequencies below 2 kHz compared to a commercially available OAE probe. The advantages and physical limitations of this system to measure OAEs in listeners with open ear canal and at low frequencies will be discussed.
Evaluating the auralization of a small room in a virtual sound environment using objective room acoustic measures

To study human auditory perception in realistic environments, loudspeaker-based reproduction techniques have recently become state-of-the-art. To evaluate the accuracy of a simulation-based room auralization of a small room, objective measures were evaluated. In particular: - early-decay time (EDT) & reverberation time (T20, T30); - clarity (C7, C50, C80); - interaural cross-correlation (IACC); - speech transmission index (STI); - direct-to-reverberant ratio (DRR). Impulse responses (IRs) were measured in an IEC listening room. The room was then modeled in the room acoustics software ODEON, and the same objective measures were evaluated for auralized versions of the playback room. The auralizations were realized using higher-order ambisonics (HOA), mixed-order ambisonics (MOA), and a nearest-loudspeaker method (NL) and reproduced in a virtual sound environment.

Exploring the Relationship Between Working Memory, Compressor Speed, and Background Noise Characteristics

Objectives: Previous work has shown that individuals with lower working memory demonstrate reduced intelligibility for speech processed with fast-acting compression amplification. This relationship has been noted in fluctuating noise, but the extent of noise modulation that must be present to elicit such an effect is unknown. This study expanded on previous study by exploring the effect of background noise modulations in relation to compression speed and working memory ability, using a range of signal to noise ratios. Design: Twenty-six older participants between ages 61 and 90 years were grouped by high or low working memory according to their performance on a reading span test. Speech intelligibility was measured for low-context sentences presented in background noise, where the noise varied in the extent of amplitude modulation. Simulated fast- or slow-acting compression amplification combined with individual frequency gain shaping was applied to compensate for the individual's hearing loss. Results: Better speech intelligibility scores were observed for participants with high working memory when fast compression was applied than when slow compression was applied. The low working memory group behaved in the opposite way and performed better under slow compression compared with fast compression. There was also a significant effect of the extent of amplitude modulation in the background noise, such that the magnitude of the score difference (fast versus slow compression) depended on the number of talkers in the background noise. The presented signal to noise ratios were not a significant factor on the measured intelligibility performance. Conclusion: In agreement with earlier research, high working memory allowed better speech intelligibility when fast compression was applied in modulated background noise. In the present experiment, that effect was present regardless of the extent of background noise modulation.
**Extériorisation sonore avec des indices auditifs et visuels discordants**

**General information**
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Technical University of Denmark
Authors: Santurette, S. (Intern), Carvajal, J. C. G. (Ekstern), Cubick, J. (Intern), Dau, T. (Intern)
Pages: 2311-2312
Publication date: 2016

**Host publication information**
Title of host publication: Proceedings of the French Acoustics Congress
Article number: CFA2016/254
Main Research Area: Technical/natural sciences
Source: PublicationPreSubmission
Source-ID: 123354184
Publication: Research - peer-review › Article in proceedings – Annual report year: 2016

**Grid - a fast threshold tracking procedure**
A new procedure, called "grid", is evaluated that allows rapid acquisition of threshold curves for psychophysics and, in particular, psychoacoustic, experiments. In this method, the parameter-response space is sampled in two dimensions within a single run. This allows the procedure to focus more experimental time investigating the vicinity of the sought-after threshold curve, compared to the current state-of-the-art methods. Therefore, time-efficiency is significantly increased and may be suitable for clinical diagnosis. While the described procedure can be used to track threshold curves in various psychoacoustic experiments, its use for measuring temporal masking curves (TMCs), based on forward masking is presented in the present study. Thresholds obtained in TMC experiments using a standard adaptive method and the new method, in a detection task, are comparable (i.e., very highly correlated).

**General information**
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Fereczkowski, M. (Intern), Dau, T. (Intern), MacDonald, E. (Intern)
Pages: 545-553
Publication date: 2016

**Host publication information**
Title of host publication: Advances in Acoustics
Publisher: Institute of Fundamental Technological Research, Polish Academy of Sciences
ISBN (Print): 978-83-65550-02-6
Main Research Area: Technical/natural sciences
Threshold-tracking, Time-efficiency, Forward masking
Electronic versions:
Grid_a_fast_threshold_tracking_procedure.pdf
Source: PublicationPreSubmission
Source-ID: 125883020
Publication: Research - peer-review › Article in proceedings – Annual report year: 2016
Impact of Background Noise and Sentence Complexity on Processing Demands during Sentence Comprehension

Speech comprehension in adverse listening conditions can be effortful even when speech is fully intelligible. Acoustical distortions typically make speech comprehension more effortful, but effort also depends on linguistic aspects of the speech signal, such as its syntactic complexity. In the present study, pupil dilations, and subjective effort ratings were recorded in 20 normal-hearing participants while performing a sentence comprehension task. The sentences were either syntactically simple (subject-first sentence structure) or complex (object-first sentence structure) and were presented in two levels of background noise both corresponding to high intelligibility. A digit span and a reading span test were used to assess individual differences in the participants’ working memory capacity (WMC). The results showed that the subjectively rated effort was mostly affected by the noise level and less by syntactic complexity. Conversely, pupil dilations increased with syntactic complexity but only showed a small effect of the noise level. Participants with higher WMC showed increased pupil responses in the higher-level noise condition but rated sentence comprehension as being less effortful compared to participants with lower WMC. Overall, the results demonstrate that pupil dilations and subjectively rated effort represent different aspects of effort. Furthermore, the results indicate that effort can vary in situations with high speech intelligibility.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Wendt, D. (Intern), Dau, T. (Intern), Hjortkjær, J. (Intern)
Number of pages: 12
Publication date: 2016
Main Research Area: Technical/natural sciences

Publication information
Journal: Frontiers in Psychology
Volume: 7
Article number: 345
ISSN (Print): 1664-1078
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 1
Scopus rating (2016): SJR 1.271 SNIP 1.006 CiteScore 2.38
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 1.28 SNIP 0.942 CiteScore 2.29
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 1.524 SNIP 1.022 CiteScore 2.71
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 1.367 SNIP 0.845 CiteScore 2.43
ISI indexed (2013): ISI indexed no
Scopus rating (2012): SJR 1.362 SNIP 0.902 CiteScore 2.39
ISI indexed (2012): ISI indexed no
Scopus rating (2011): SJR 0.726 SNIP 0.422
ISI indexed (2011): ISI indexed no
Original language: English
Effort, Processing demands, Pupillometry, Syntactic complexity, Background noise, Working memory capacity, Reading span, Digit span
Electronic versions:
Wendt_etal_2016.pdf
DOIs:
10.3389/fpsyg.2016.00345
Source: PublicationPreSubmission
Source-ID: 122180170
Publication: Research - peer-review › Journal article – Annual report year: 2016
Individual Hearing Loss: Characterization, Modelling, Compensation Strategies

It is well-established that hearing loss does not only lead to a reduction of hearing sensitivity. Large individual differences are typically observed among listeners with hearing impairment in a wide range of suprathreshold auditory measures. In many cases, audiometric thresholds cannot fully account for such individual differences, which make it challenging to find adequate compensation strategies in hearing devices. How to characterize, model, and compensate for individual hearing loss were the main topics of the fifth International Symposium on Auditory and Audiological Research (ISAAR), held in Nyborg, Denmark, in August 2015. The following collection of papers results from some of the work that was presented and discussed at the symposium.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Copenhagen, University of Southern Denmark
Authors: Santurette, S. (Intern), Dau, T. (Intern), Christensen-Dalsgaard, J. (Ekstern), Tranebjærg, L. (Ekstern), Andersen, T. (Ekstern), Poulsen, T. (Intern)
Number of pages: 2
Publication date: 2016
Conference: 5th International Symposium on Auditory and Audiological Research, Nyborg, Denmark, 26/08/2015 - 26/08/2015
Main Research Area: Technical/natural sciences

Publication information
Journal: Trends in Hearing
Volume: 20
Issue number: Special Issue
ISSN (Print): 2331-2165
Ratings:
Web of Science (2018): Indexed yes
Web of Science (2017): Indexed Yes
Scopus rating (2016): CiteScore 3.61
Web of Science (2016): Indexed yes
Scopus rating (2015): SJR 1.608 SNIP 1.468 CiteScore 2.21
Web of Science (2015): Indexed yes
Scopus rating (2014): SJR 1.006 SNIP 1.099 CiteScore 0
Web of Science (2014): Indexed yes
Scopus rating (2013): SJR 0.953 SNIP 1.009
Scopus rating (2012): SJR 0.981 SNIP 0.909
Scopus rating (2011): SJR 1.164 SNIP 1.551
Scopus rating (2010): SJR 1.612 SNIP 1.997
Scopus rating (2009): SJR 1.263 SNIP 1.208
Scopus rating (2008): SJR 1.87 SNIP 1.678
Scopus rating (2007): SJR 2.255 SNIP 2.124
Scopus rating (2006): SJR 0.97 SNIP 1.095
Scopus rating (2005): SJR 0.892 SNIP 0.572
Scopus rating (2004): SJR 0.289 SNIP 0.259
Scopus rating (2003): SJR 0.524 SNIP 1.142
Scopus rating (2002): SJR 0.377 SNIP 1.134
Scopus rating (2001): SJR 0.308 SNIP 1.09
Scopus rating (2000): SJR 0.133 SNIP 0
Scopus rating (1999): SJR 0.297 SNIP 0.706
Original language: English
Individual differences, Hearing loss characterization, Auditory modelling, Hearing aids, Cochlear implants
Electronic versions:
Trends_in_Hearing_2016_Santurette_2331216516655890.pdf
DOI:
10.1177/2331216516655890

Bibliographical note
Creative Commons CC-BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 3.0 License (http://www.creativecommons.org/licenses/by-nc/3.0/) which permits non-commercial use,
Investigating low-frequency compression using the Grid method

There is an ongoing discussion about whether the amount of cochlear compression in humans at low frequencies (below 1 kHz) is as high as that at higher frequencies. It is controversial whether the compression affects the slope of the off-frequency forward masking curves at those frequencies.

Here, the Grid method with a 2-interval 1-up 3-down tracking rule was applied to estimate forward masking curves at two characteristic frequencies: 500 Hz and 4000 Hz. The resulting curves and the corresponding basilar membrane input-output (BM I/O) functions were found to be comparable to those reported in literature. Moreover, slopes of the low-level portions of the BM I/O functions estimated at 500 Hz were examined, to determine whether the 500-Hz off-frequency forward masking curves were affected by compression. Overall, the collected data showed a trend confirming the compressive behaviour. However, the analysis was complicated by unexpectedly steep portions of the collected on- and off-frequency forward masking curves.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Fereczkowski, M. (Intern), Dau, T. (Intern), MacDonald, E. (Intern)
Pages: 413-420
Publication date: 2016

Host publication information
Title of host publication: Proceedings of ISAAR 2015
Editors: Santurette, S., Dau, T., Dalsgaard, J. C., Tranebjaerg, L., Andersen, T.
ISBN (Print): 978-87-990013-5-4
Main Research Area: Technical/natural sciences
Conference: 5th International Symposium on Auditory and Audiological Research, Nyborg, Denmark, 26/08/2015 - 26/08/2015
Electronic versions:
Links:
Source: PublicationPreSubmission
Source-ID: 125379555
Publication: Research - peer-review › Article in proceedings – Annual report year: 2016

Lateralized speech perception in normal-hearing and hearing-impaired listeners and its relationship to temporal processing

This study investigated the role of temporal fine structure (TFS) coding in spatially complex, lateralized listening tasks. Speech reception thresholds (SRTs) were measured in young normal-hearing (NH) and two groups of elderly hearing-impaired (HI) listeners in the presence of speech-shaped noise and different interfering talker conditions. The HI subjects had either a mild or moderate hearing loss above 1.5 kHz and reduced audibility was compensated for individually in the speech tests. The target and masker streams were presented as coming from the same or from the opposite side of the head by introducing 0.7-ms interaural time differences (ITD) between the ears. To assess the robustness of TFS coding, frequency discrimination thresholds (FDTs) and interaural phase difference thresholds (IPDTs) were measured at 250 Hz. While SRTs of the NH subjects were clearly better than those of the HI listeners, group differences in binaural benefit due to spatial separation of the maskers from the target remained small. Neither the FDT nor the IPDT tasks showed a clear correlation pattern with the SRTs or with the amount of binaural benefit, respectively. The results suggest that, although HI listeners with normal hearing in the low-frequency range might have elevated SRTs, the binaural benefit they experience due to spatial separation of competing sources can remain similar to that of NH listeners.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Eriksholm Research Centre
Authors: Locsei, G. (Intern), Pedersen, J. H. (Ekstern), Laugesen, S. (Ekstern), Santurette, S. (Intern), Dau, T. (Intern), MacDonald, E. (Intern)
Pages: 389-396
Publication date: 2016

Host publication information
Model-based fitting of compression settings using narrowband stimuli

Most state-of-the-art hearing aids apply multi-channel dynamic-range compression (DRC). Studies using speech intelligibility as an outcome measure have shown mixed results in terms of the benefits of compression over linear amplification (e.g., Davies-Venn et al. 2009; Goedegebure et al. 2001, 2002; Kates 2010; Olsen et al. 2005; Souza et al. 1999, 2012; Yund and Buckles 1995a,b). Compression provides increased audibility of speech components, but at the same time introduces distortion of spectral and temporal envelopes of speech. The two effects may offset each other, depending on what cues the individual hearing-impaired listeners rely on. Therefore, it is difficult to disentangle them when speech recognition is used as an outcome measure. Edwards (2002) suggested using a set of relatively simple outcome measures, based on narrowband signals, for the evaluation of hearing-aid signal processing. We present a compression design that has been optimized, within the framework of a computational model, for improving the performance of (aided) hearing impaired listeners in temporal and spectral resolution-related tasks.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Sonova U.S. Corporate Services
Authors: Kowalewski, B. (Intern), Fereczkowski, M. (Intern), MacDonald, E. (Intern), Strelcyk, O. (Ekstern), Dau, T. (Intern)
Number of pages: 1
Publication date: 2016
Event: Poster session presented at International Hearing Aid Conference 2016, Tahoe, CA, United States.
Main Research Area: Technical/natural sciences
Electronic versions:
IHCON2016_Borys_verFinal.pdf
Source: PublicationPreSubmission
Source-ID: 125882820
Publication: Research › Poster – Annual report year: 2016

Modeling spectro-temporal modulation perception in normal hearing listeners

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Sanchez Lopez, R. (Intern), Dau, T. (Intern)
Number of pages: 12
Publication date: 2016

Host publication information
Title of host publication: Proceedings of Inter-Noise 2016
Publisher: Deutsche Gesellschaft für Akustik
Editor: Kropp, W.
ISBN (Electronic): 978-3-939296-11-9
BFI conference series: Inter-Noise (5010071)
Main Research Area: Technical/natural sciences
Conference: 45th International Congress and Exposition on Noise Control Engineering, INTER-NOISE 2016, Hamburg, Germany, 21/08/2016 - 21/08/2016
Spectro-temporal modulation, Auditory modeling
Electronic versions:
pdfdoc.pdf
Source: PublicationPreSubmission
Source-ID: 124363169
Publication: Research - peer-review › Article in proceedings – Annual report year: 2016
Modelling the Perceptual Components of Loudspeaker Distortion

While non-linear distortion in loudspeakers decreases audio quality, the perceptual consequences can vary substantially. This paper investigates the metric $R_{\text{nonlin}}$ [1] which was developed to predict subjective measurements of sound quality in nonlinear systems. The generalisability of the metric in a practical setting was explored across a range of different loudspeakers and signals. Overall, the correlation of $R_{\text{nonlin}}$ predictions with subjective ratings was poor. Based on further investigation, an additional normalization step is proposed, which substantially improves the ability of $R_{\text{nonlin}}$ to predict the perceptual consequences of non-linear distortion.

Objective and Perceptual Evaluation of a Virtual Sound Environment System.

Hearing aid (HA) users often have difficulties following a conversation in challenging listening situations involving multiple talkers, background noise, and reverberation. To improve their listening performance, the algorithms in modern HAs have become increasingly complex. Yet, most HA testing is still performed in unrealistically simple setups.

Publication: Research - peer-review › Conference abstract in journal – Annual report year: 2017

**Modelling the Perceptual Components of Loudspeaker Distortion**

While non-linear distortion in loudspeakers decreases audio quality, the perceptual consequences can vary substantially. This paper investigates the metric $R_{\text{nonlin}}$ [1] which was developed to predict subjective measurements of sound quality in nonlinear systems. The generalisability of the metric in a practical setting was explored across a range of different loudspeakers and signals. Overall, the correlation of $R_{\text{nonlin}}$ predictions with subjective ratings was poor. Based on further investigation, an additional normalization step is proposed, which substantially improves the ability of $R_{\text{nonlin}}$ to predict the perceptual consequences of non-linear distortion.

**Objective and Perceptual Evaluation of a Virtual Sound Environment System.**

Hearing aid (HA) users often have difficulties following a conversation in challenging listening situations involving multiple talkers, background noise, and reverberation. To improve their listening performance, the algorithms in modern HAs have become increasingly complex. Yet, most HA testing is still performed in unrealistically simple setups.
Objective assessment of stream segregation abilities of CI users as a function of electrode separation

Auditory streaming is a perceptual process by which the human auditory system organizes sounds from different sources into perceptually meaningful elements. Segregation of sound sources is important, among others, for understanding speech in noisy environments, which is especially challenging for cochlear implant (CI) users. Despite its high relevance in many daily situations, the number of studies investigating segregation abilities of CI listeners is limited and their findings are contradictory (e.g. Cooper and Roberts, 2009; Marozeau et al, 2013). Moreover, while most of the previous research assessed obligatory stream segregation, little attention has been given to voluntary stream segregation, a process where the listener actively tries to segregate the sounds. It is therefore unclear whether CI users are able to experience voluntary stream segregation as a function of electrode separation and whether this is perceived to occur instantaneously or to build-up over time.

On-site and laboratory soundscape evaluations of three recreational urban spaces

Soundscape quality was evaluated using four subjective psychological rating factors in three recreational urban spaces in which water and a variation of other natural and anthropogenic sound sources were present. The noise level was measured at each site during occupant peak flows and recordings for listening experiments were made simultaneously. Listeners answered questionnaires either on site or following playback of the recordings in the laboratory, with or without access to each site’s visual context. They rated their perception of loudness, acceptance, stressfulness, and comfort, along with their preference toward eight sound sources. The comfort ratings were negatively correlated with loudness and stressfulness and positively correlated with acceptance. The sound level was found to be a good predictor of these subjective parameters in the laboratory, but not on site. Moreover, the availability of the visual context in the listening experiment had no effect on the ratings. The presence of trees and water was also found to increase on-site comfort. Generally, the participants were more positive towards natural sound sources on-site. Overall, the results suggest that on-site context plays an important role for evaluating acoustic comfort in urban recreational areas.
Outcome measures based on classification performance fail to predict the intelligibility of binary-masked speech

To date, the most commonly used outcome measure for assessing ideal binary mask estimation algorithms is based on the difference between the hit rate and the false alarm rate (H-FA). Recently, the error distribution has been shown to substantially affect intelligibility. However, H-FA treats each mask unit independently and does not take into account how errors are distributed. Alternatively, algorithms can be evaluated with the short-time objective intelligibility (STOI) metric using the reconstructed speech. This study investigates the ability of H-FA and STOI to predict intelligibility for binary-masked speech using masks with different error distributions. The results demonstrate the inability of H-FA to predict the behavioral intelligibility and also illustrate the limitations of STOI. Since every estimation algorithm will make errors that are distributed in different ways, performance evaluations should not be made solely on the basis of these metrics.
Perceptual Spaces Induced by Cochlear Implant All-Polar Stimulation Mode

It has been argued that a main limitation of the cochlear implant is the spread of current induced by each electrode, which activates an inappropriately large range of sensory neurons. To reduce this spread, an alternative stimulation mode, the all-polar mode, was tested with five participants. It was designed to activate all the electrodes simultaneously with appropriate current levels and polarities to recruit narrower regions of auditory nerves at specific intracochlear electrode positions (denoted all-polar electrodes). In this study, the all-polar mode was compared with the current commercial stimulation mode: the monopolar mode. The participants were asked to judge the sound dissimilarity between pairs of two-electrode pulse-train stimuli that differed in the electrode positions and were presented in either monopolar or all-polar mode with pulses on the two electrodes presented either sequentially or simultaneously. The dissimilarity ratings were analyzed using a multidimensional scaling technique and three-dimensional stimulus perceptual spaces were produced. For all the conditions (mode and simultaneity), the first perceptual dimension was highly correlated with the position of the most apical activated electrode of the electrical stimulation and the second dimension with the position of the most basal electrode. In both sequential and simultaneous conditions, the monopolar and all-polar stimuli were significantly separated by a third dimension, which may indicate that all-polar stimuli have a perceptual quality that differs from monopolar stimuli. Overall, the results suggest that both modes might successfully represent spectral information in a sound processing strategy.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Marozeau, J. (Intern), McKay, C. M. (Ekstern)
Publication date: 2016
Pitch Discrimination in Musicians and Non-Musicians: Effects of Harmonic Resolvability and Processing Effort

Musicians typically show enhanced pitch discrimination abilities compared to non-musicians. The present study investigated this perceptual enhancement behaviorally and objectively for resolved and unresolved complex tones to clarify whether the enhanced performance in musicians can be ascribed to increased peripheral frequency selectivity and/or to a different processing effort in performing the task. In a first experiment, pitch discrimination thresholds were obtained for harmonic complex tones with fundamental frequencies (F0s) between 100 and 500 Hz, filtered in either a low- or a high-frequency region, leading to variations in the resolvability of audible harmonics. The results showed that pitch discrimination performance in musicians was enhanced for resolved and unresolved complexes to a similar extent. Additionally, the harmonics became resolved at a similar F0 in musicians and non-musicians, suggesting similar peripheral frequency selectivity in the two groups of listeners. In a follow-up experiment, listeners' pupil dilations were measured as an indicator of the required effort in performing the same pitch discrimination task for conditions of varying resolvability and task difficulty. Pupillometry responses indicated a lower processing effort in the musicians versus the non-musicians,
although the processing demand imposed by the pitch discrimination task was individually adjusted according to the behavioral thresholds. Overall, these findings indicate that the enhanced pitch discrimination abilities in musicians are unlikely to be related to higher peripheral frequency selectivity and may suggest an enhanced pitch representation at more central stages of the auditory system in musically trained listeners.

**General information**

*State:* Published  
*Organisations:* Department of Electrical Engineering, Hearing Systems  
*Authors:* Bianchi, F. (Intern), Santurette, S. (Intern), Wendt, D. (Intern), Dau, T. (Intern)  
*Number of pages:* 11  
*Pages:* 69-79  
*Publication date:* 2016  
*Main Research Area:* Technical/natural sciences

**Publication information**

*Journal:* J A R O  
*Volume:* 17  
*Issue number:* 1  
*ISSN (Print):* 1525-3961  
*Ratings:*  
*BFI (2018):* BFI-level 2  
*Web of Science (2018):* Indexed yes  
*BFI (2017):* BFI-level 2  
*Web of Science (2017):* Indexed yes  
*BFI (2016):* BFI-level 2  
*Web of Science (2016):* Indexed yes  
*BFI (2015):* BFI-level 2  
*Scopus rating (2015):* SJR 1.688 SNIP 1.48 CiteScore 2.95  
*Web of Science (2015):* Indexed yes  
*BFI (2014):* BFI-level 2  
*Scopus rating (2014):* SJR 1.592 SNIP 1.371 CiteScore 2.84  
*BFI (2013):* BFI-level 2  
*Scopus rating (2013):* SJR 1.333 SNIP 1.432 CiteScore 2.67  
*ISI indexed (2013):* ISI indexed yes  
*Web of Science (2013):* Indexed yes  
*BFI (2012):* BFI-level 2  
*Scopus rating (2012):* SJR 1.535 SNIP 1.299 CiteScore 2.74  
*ISI indexed (2012):* ISI indexed yes  
*BFI (2011):* BFI-level 2  
*Scopus rating (2011):* SJR 1.416 SNIP 1.5 CiteScore 2.98  
*ISI indexed (2011):* ISI indexed yes  
*BFI (2010):* BFI-level 2  
*Scopus rating (2010):* SJR 1.617 SNIP 1.406  
*Web of Science (2010):* Indexed yes  
*BFI (2009):* BFI-level 2  
*Scopus rating (2009):* SJR 1.406 SNIP 1.17  
*BFI (2008):* BFI-level 2  
*Scopus rating (2008):* SJR 1.456 SNIP 1.066  
*Scopus rating (2007):* SJR 1.418 SNIP 1.123  
*Scopus rating (2006):* SJR 1.398 SNIP 1.074  
*Scopus rating (2005):* SJR 1.463 SNIP 1.065  
*Scopus rating (2004):* SJR 1.283 SNIP 1.221  
*Scopus rating (2003):* SJR 1.222 SNIP 1.038  
*Scopus rating (2002):* SJR 1 SNIP 0.971  
*Scopus rating (2001):* SJR 0.637 SNIP 0.469  
*Original language:* English  
*Pitch discrimination, Resolved complex tones, Unresolved complex tones, Musicians, Pupillometry, Processing effort*
Predicting binaural speech intelligibility using the signal-to-noise ratio in the envelope power spectrum domain

This study proposes a binaural extension to the multi-resolution speech-based envelope power spectrum model (mr-sEPSM) [Jørgensen, Ewert, and Dau (2013). J. Acoust. Soc. Am. 134, 436–446]. It consists of a combination of better-ear (BE) and binaural unmasking processes, implemented as two monaural realizations of the mr-sEPSM combined with a short-term equalization-cancellation process, and uses the signal-to-noise ratio in the envelope domain (SNRenv) as the decision metric. The model requires only two parameters to be fitted per speech material and does not require an explicit frequency weighting. The model was validated against three data sets from the literature, which covered the following effects: the number of maskers, the masker types [speech-shaped noise (SSN), speech-modulated SSN, babble, and reversed speech], the masker(s) azimuths, reverberation on the target and masker, and the interaural time difference of the target and masker. The Pearson correlation coefficient between the simulated speech reception thresholds and the data across all experiments was 0.91. A model version that considered only BE processing performed similarly (correlation coefficient of 0.86) to the complete model, suggesting that BE processing could be considered sufficient to predict intelligibility in most realistic conditions.
Predicting Detectability and Annoyance of EV Warning Sounds using Partial Loudness

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Technical University of Denmark, Bruel and Kjaer Sound and Vibration Measurement A/S, Korea Advanced Institute of Science & Technology
Authors: Jacobsen, G. N. (Intern), Ih, J. (Ekstern), Song, W. (Ekstern), MacDonald, E. (Intern)
Pages: 1706-1715
Publication date: 2016

Host publication information
Title of host publication: Proceedings of the 45th International Congress and Exposition on Noise Control Engineering
Publisher: Deutsche Gesellschaft für Akustik
Editor: Kropp, W.
Predicting effects of non-linear frequency compression and impulse-noise suppression on consonant perception

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Sonova AG
Authors: Zaaïr, J. (Intern), Schmitt, N. (Ekstern), Derleth, R. (Ekstern), Dau, T. (Intern)
Number of pages: 1
Publication date: 2016
Event: Poster session presented at International Hearing Aid Conference 2016, Tahoe, CA, United States.
Main Research Area: Technical/natural sciences
Electronic versions:
IHCON_poster_Zaaïr_et_al.pdf
Source: PublicationPreSubmission
Source-ID: 125852134
Publication: Research - peer-review › Poster – Annual report year: 2016

Predicting masking release of lateralized speech
Locsei et al. (2015) [Speech in Noise Workshop, Copenhagen, 46] measured speech reception thresholds (SRTs) in anechoic conditions where the target speech and the maskers were lateralized using interaural time delays. The maskers were speech-shaped noise (SSN) and reversed babble with 2, 4, or 8 talkers. For a given interferer type, the number of maskers presented on the target’s side was varied, such that none, some, or all maskers were presented on the same side as the target. In general, SRTs did not vary significantly when at least one masker was presented on the same side as the target. The largest masking release (MR) was observed when all maskers were on the opposite side of the target. The data in the conditions containing only energetic masking and modulation masking could be accounted for using a binaural extension of the speech-based envelope power spectrum model [sEPSM; Jørgensen et al., 2013, J. Acoust. Soc. Am. 130], which uses a short-term equalization-cancellation process to model binaural unmasking. In the conditions where informational masking (IM) was involved, the predicted SRTs were lower than the measured values because the model is blind to confusions experienced by the listeners. Additional simulations suggest that, in these conditions, it would be possible to estimate the confusions, and thus the amount of IM, based on the similarity of the target and masker representations in the envelope power domain.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Chabot-Leclerc, A. (Intern), MacDonald, E. (Intern), Dau, T. (Intern)
Number of pages: 8
Publication date: 2016

Host publication Information
Title of host publication: Proceedings of ISAAR 2015
Editors: Santurette, S., Dau, T., Dalsgaard, J. C., Tranebjærg, T., Andersen, T.
ISBN (Print): 978-87-990013-5-4
Main Research Area: Technical/natural sciences
Conference: 5th International Symposium on Auditory and Audiological Research, Nyborg, Denmark, 26/08/2015 - 26/08/2015
Source: PublicationPreSubmission
Source-ID: 123036868
Publication: Research - peer-review › Article in proceedings – Annual report year: 2016

Predicting speech intelligibility based on a correlation metric in the envelope power spectrum domain
model, named sEPSMcorr, is shown to account for the effects of stationary and fluctuating additive interferers as well as for the effects of non-linear distortions, such as spectral subtraction, phase jitter, and ideal time frequency segregation (ITFS). The model shows a broader predictive range than both the original mr-sEPSM (which fails in the phase-jitter and ITFS conditions) and STOI (which fails to predict the influence of fluctuating interferers), albeit with lower accuracy than the source models in some individual conditions. Similar to other models that employ a short-term correlation-based back end, including STOI, the proposed model fails to account for the effects of room reverberation on speech intelligibility.

Overall, the model might be valuable for evaluating the effects of a large range of interferers and distortions on speech intelligibility, including consequences of hearing impairment and hearing-instrument signal processing.

**General information**

State: Published  
Organisations: Department of Electrical Engineering, Hearing Systems  
Authors: Relaño-Iborra, H. (Intern), May, T. (Intern), Zaar, J. (Intern), Scheidiger, C. (Intern), Dau, T. (Intern)  
Pages: 2670–2679  
Publication date: 2016  
Main Research Area: Technical/natural sciences

**Publication information**

Journal: Journal of the Acoustical Society of America  
Volume: 140  
Issue number: 4  
ISSN (Print): 0001-4966  
Ratings:  
BFI (2018): BFI-level 2  
Web of Science (2018): Indexed yes  
BFI (2017): BFI-level 2  
Web of Science (2017): Indexed yes  
BFI (2016): BFI-level 2  
Web of Science (2016): Indexed yes  
BFI (2015): BFI-level 2  
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27  
Web of Science (2016): Indexed yes  
BFI (2014): BFI-level 2  
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77  
Web of Science (2015): Indexed yes  
BFI (2013): BFI-level 2  
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8  
Web of Science (2014): Indexed yes  
BFI (2012): BFI-level 2  
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2  
ISI indexed (2013): ISI indexed yes  
Web of Science (2013): Indexed yes  
BFI (2012): BFI-level 2  
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75  
ISI indexed (2012): ISI indexed yes  
Web of Science (2012): Indexed yes  
BFI (2011): BFI-level 2  
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68  
ISI indexed (2011): ISI indexed yes  
Web of Science (2011): Indexed yes  
BFI (2010): BFI-level 2  
Scopus rating (2010): SJR 0.754 SNIP 1.528  
Web of Science (2010): Indexed yes  
BFI (2009): BFI-level 2  
Scopus rating (2009): SJR 0.783 SNIP 1.717  
Web of Science (2009): Indexed yes  
BFI (2008): BFI-level 2  
Scopus rating (2008): SJR 0.848 SNIP 1.633  
Web of Science (2008): Indexed yes  
Scopus rating (2007): SJR 0.865 SNIP 1.647
Simultaneous measurement of auditory-steady-state responses and otoacoustic emissions to estimate peripheral compression

Assessment of the compressive nonlinearity in the hearing system provides useful information about the inner ear. Auditory-steady state responses (ASSR) have recently been used to estimate the state of the compressive nonlinearity in the peripheral auditory system. Since it is commonly assumed that outer hair cells in the inner ear play an important role in the compressive nonlinearity, it is desirable to selectively obtain information about the inner ear. In the current study, the signal in the ear canal present during ASSR measurements is utilized to extract sinusoidally-amplitude modulated otoacoustic emissions (SAMOAEs). It is hypothesized that the stimulus used to evoke ASSRs will cause acoustic energy to be reflected back from the inner ear into the ear canal, where it can be picked up as an otoacoustic emission (OAE) and provide information about cochlear processing. Results indicate that SAMOAEs can be extracted while measuring ASSRs using sinusoidally-amplitude modulated tones. However, comparison of simulations using a transmission model and the data show that the SAMOAE measured above 50 dB SPL are strongly influenced by the system distortion. A robust extraction and evaluation of SAMOAE in connection with ASSR may be possible by a proposed method to minimize the distortion. The ability to evaluate SAMOAE over a large input level range during ASSR measurement will provide information about the state of the peripheral auditory system without the need of additional measurement time.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Sanchez Lopez, R. (Intern), Epp, B. (Intern)
Number of pages: 8
Publication date: 2016

Host publication information
Title of host publication: Proceedings of ISAAR 2015
Editors: Santurette, S., Dau, T., Dalsgaard, J. C., Tranebjaerg, L., Andersen, T.
ISBN (Print): 978-87-990013-5-4
Main Research Area: Technical/natural sciences
Conference: 5th International Symposium on Auditory and Audiological Research, Nyborg, Denmark, 26/08/2015 - 26/08/2015
Source: PublicationPreSubmission
Sources of Variability in Consonant Perception and Implications for Speech Perception Modeling

The present study investigated the influence of various sources of response variability in consonant perception. A distinction was made between source-induced variability and receiver-related variability. The former refers to perceptual differences induced by differences in the speech tokens and/or the masking noise tokens; the latter describes perceptual differences caused by within- and across-listener uncertainty. Consonant-vowel combinations (CVs) were presented to normal-hearing listeners in white noise at six different signal-to-noise ratios. The obtained responses were analyzed with respect to the considered sources of variability using a measure of the perceptual distance between responses. The largest effect was found across different CVs. For stimuli of the same phonetic identity, the speech-induced variability across and within talkers and the across-listener variability were substantial and of similar magnitude. Even time-shifts in the waveforms of white masking noise produced a significant effect, which was well above the within-listener variability (the smallest effect). Two auditory-inspired models in combination with a template-matching back end were considered to predict the perceptual data. In particular, an energy-based and a modulation-based approach were compared. The suitability of the two models was evaluated with respect to the source-induced perceptual distance and in terms of consonant recognition rates and consonant confusions. Both models captured the source-induced perceptual distance remarkably well. However, the modulation-based approach showed a better agreement with the data in terms of consonant recognition and confusions. The results indicate that low-frequency modulations up to 16 Hz play a crucial role in consonant perception.

Spatial Hearing with Incongruent Visual or Auditory Room Cues

In day-to-day life, humans usually perceive the location of sound sources as outside their heads. This externalized auditory spatial perception can be reproduced through headphones by recreating the sound pressure generated by the source at the listener’s eardrums. This requires the acoustical features of the recording environment and listener’s anatomy to be recorded at the listener’s ear canals. Although the resulting auditory images can be indistinguishable from real-world sources, their externalization may be less robust when the playback and recording environments differ. Here we tested whether a mismatch between playback and recording room reduces perceived distance, azimuthal direction, and compactness of the auditory image, and whether this is mostly due to incongruent auditory cues or to expectations.
generated from the visual impression of the room. Perceived distance ratings decreased significantly when collected in a more reverberant environment than the recording room, whereas azimuthal direction and compactness remained room independent. Moreover, modifying visual room-related cues had no effect on these three attributes, while incongruent auditory room-related cues between the recording and playback room did affect distance perception. Consequently, the external perception of virtual sounds depends on the degree of congruency between the acoustical features of the environment and the stimuli.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems, Department of Electrical Engineering, Hearing Systems
Authors: Gil Carvajal, J. C. (Intern), Cubick, J. (Intern), Santurette, S. (Intern), Dau, T. (Intern)
Number of pages: 10
Publication date: 2016
Main Research Area: Technical/natural sciences

Publication information
Journal: Scientific Reports
Volume: 6
Article number: 37342
ISSN (Print): 2045-2322
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 4.63 SJR 1.625 SNIP 1.401
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 2.057 SNIP 1.684 CiteScore 5.3
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 2.103 SNIP 1.544 CiteScore 4.75
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 1.886 SNIP 1.51 CiteScore 4.06
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): SJR 1.458 SNIP 0.896 CiteScore 2.44
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
ISI indexed (2011): ISI indexed no
Original language: English
Electronic versions:
srep37342.pdf
DOIs:
10.1038/srep37342
Links:
http://www.nature.com/articles/srep37342
Publication: Research - peer-review › Journal article – Annual report year: 2016

Spectral and temporal cues for perception of material and action categories in impacted sound sources
In two experiments, similarity ratings and categorization performance with recorded impact sounds representing three material categories (wood, metal, glass) being manipulated by three different categories of action (drop, strike, rattle) were examined. Previous research focusing on single impact sounds suggests that temporal cues related to damping are essential for material discrimination, but spectral cues are potentially more efficient for discriminating materials manipulated by different actions that include multiple impacts (e.g., dropping, rattling). Perceived similarity between material
categories across different actions was correlated with the distribution of long-term spectral energy (spectral centroid). Similarity between action categories was described by the temporal distribution of envelope energy (temporal centroid) or by the density of impacts. Moreover, perceptual similarity correlated with the pattern of confusion in categorization judgments. Listeners tended to confuse materials with similar spectral centroids, and actions with similar temporal centroids and onset densities. To confirm the influence of these different features, spectral cues were removed by applying the envelopes of the original sounds to a broadband noise carrier. Without spectral cues, listeners retained sensitivity to action categories but not to material categories. Conversely, listeners recognized material but not action categories after envelope scrambling that preserved long-term spectral content.
Spectro-temporal modulation sensitivity and discrimination in normal hearing and hearing-impaired listeners

When a signal varies in its properties along the time and frequency, this is considered a modulation. Speech signals exhibit temporal and spectral modulations. The sensitivity to these modulations has been studied in normal-hearing (NH) listeners, yielding temporal, spectral and spectro-temporal modulation transfer functions (Dau et al. 1997, Eddins & Bero 2007, Chi et al. 1999). Recently, Mehraei et al. (2014) showed significant differences between normal-hearing and hearing-impaired (HI) listeners in spectro-temporal modulation (STM) detection and also the relation between STM sensitivity to speech intelligibility in noise. Moreover, Henry et al. (2005) showed large differences in STM discriminations tasks. The present study attempted to establish the limits of STM perception in NH listeners and two groups of HI (with either good or poor speech intelligibility).

Subjective evaluation of restaurant acoustics in a virtual sound environment

Many restaurants have smooth rigid surfaces made of wood, steel, glass, and concrete. This often results in a lack of sound absorption. Such restaurants are notorious for high sound noise levels during service that most owners actually desire for representing vibrant eating environments, although surveys report that noise complaints are on par with poor service. This study investigated the relation between objective acoustic parameters and subjective evaluation of acoustic comfort at five restaurants in terms of three parameters: noise annoyance, speech intelligibility, and privacy. At each
location, customers filled out questionnaire surveys, acoustic parameters were measured, and recordings of restaurant acoustic scenes were obtained with a 64-channel spherical array. The acoustic scenes were reproduced in a virtual sound environment (VSE) with 64 loudspeakers placed in an anechoic room, where listeners performed subjective evaluation of noise annoyance and privacy and a speech intelligibility test for each restaurant noise background. It was found that subjective evaluations of acoustic comfort correlate with occupancy rates and measured noise levels, that survey and listening test results agreed well and that, in the VSE, speech reception thresholds were similar for the five reproduced restaurant backgrounds.

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Acoustic Technology
Authors: Nielsen, N. Ø. (Intern), Marschall, M. (Intern), Santurette, S. (Intern), Jeong, C. (Intern)
Pages: 6140-6149
Publication date: 2016

**Host publication information**

Title of host publication: Proceedings of 45th International Congress and Exposition on Noise Control Engineering
Publisher: German Acoustical Society (DEGA)
Editor: Kropp, W.
ISBN (Electronic): 978-3-939296-11-9
BFI conference series: Inter-Noise (5010071)
Main Research Area: Technical/natural sciences
Conference: 45th International Congress and Exposition on Noise Control Engineering, INTER-NOISE 2016, Hamburg, Germany, 21/08/2016 - 21/08/2016
Restaurant noise, Virtual sound reproduction, Noise annoyance
Source: PublicationPreSubmission
Source-ID: 125556153
Publication: Research - peer-review › Article in proceedings – Annual report year: 2016

**Temporal Fine-Structure Coding and Lateralized Speech Perception in Normal-Hearing and Hearing-Impaired Listeners**

This study investigated the relationship between speech perception performance in spatially complex, lateralized listening scenarios and temporal fine-structure (TFS) coding at low frequencies. Young normal-hearing (NH) and two groups of elderly hearing-impaired (HI) listeners with mild or moderate hearing loss above 1.5kHz participated in the study. Speech reception thresholds (SRTs) were estimated in the presence of either speech-shaped noise, two-, four-, or eight-talker babble played reversed, or a nonreversed two-talker masker. Target audibility was ensured by applying individualized linear gains to the stimuli, which were presented over headphones. The target and masker streams were lateralized to the same or to opposite sides of the head by introducing 0.7-ms interaural time differences between the ears. TFS coding was assessed by measuring frequency discrimination thresholds and interaural phase difference thresholds at 250Hz. NH listeners had clearly better SRTs than the HI listeners. However, when maskers were spatially separated from the target, the amount of SRT benefit due to binaural unmasking differed only slightly between the groups. Neither the frequency discrimination threshold nor the interaural phase difference threshold tasks showed a correlation with the SRTs or with the amount of masking release due to binaural unmasking, respectively. The results suggest that, although HI listeners with normal hearing thresholds below 1.5kHz experienced difficulties with speech understanding in spatially complex environments, these limitations were unrelated to TFS coding abilities and were only weakly associated with a reduction in binaural-unmasking benefit for spatially separated competing sources.

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Department of Acoustic Technology, Eriksholm Research Centre
Authors: Locsei, G. (Intern), Pedersen, J. H. (Ekstern), Laugesen, S. (Intern), Santurette, S. (Intern), Dau, T. (Intern), MacDonald, E. (Intern)
Number of pages: 15
Publication date: 2016
Main Research Area: Technical/natural sciences

**Publication information**

Journal: Trends in Hearing
Volume: 20
Issue number: 0
ISSN (Print): 2331-2165
Ratings:
Web of Science (2018): Indexed yes
Web of Science (2017): Indexed Yes
The Effect of Tactile Cues on Auditory Stream Segregation Ability of Musicians and Nonmusicians.

Difficulty perceiving music is often cited as one of the main problems facing hearing-impaired listeners. It has been suggested that musical enjoyment could be enhanced if sound information absent due to impairment is transmitted via other sensory modalities such as vision or touch. In this study, we test whether tactile cues can be used to segregate 2 interleaved melodies. Twelve musicians and 12 nonmusicians were asked to detect changes in a 4-note repeated melody interleaved with a random melody. In order to perform this task, the listener must be able to segregate the target melody from the random melody. Tactile cues were applied to the listener’s fingers on half of the blocks. Results showed that tactile cues can significantly improve the melodic segregation ability in both musician and nonmusician groups in challenging listening conditions. Overall, the musician group performance was always better; however, the magnitude of improvement with the introduction of tactile cues was similar in both groups. This study suggests that hearing-impaired listeners could potentially benefit from a system transmitting such information via a tactile modality.
The Influence of Visual Cues on Sound Externalization

Background: The externalization of virtual sounds reproduced via binaural headphone-based auralization systems has been reported to be less robust when the listening environment differs from the room in which binaural room impulse responses (BRIRs) were recorded. It has been debated whether this is due to incongruent auditory cues between the recording and playback room during sound reproduction or to an expectation effect from the visual impression of the room. This study investigated the influence of a priori acoustic and visual knowledge of the playback room on sound externalization.

Methods: Eighteen naive listeners rated the externalization of virtual stimuli in terms of perceived distance, azimuthal localization, and compactness in three rooms: 1) a standard IEC listening room, 2) a small reverberant room, and 3) a large dry room. Before testing, individual BRIRs were recorded in room 1 while listeners wore both earplugs and blindfolds. Half of the listeners were then blindfolded during testing but were provided auditory awareness of the room via a controlled noise source (condition A). The other half could see the room but were shielded from room-related acoustic input and tested without the controlled noise source (condition V). All listeners were also tested with all cues available (condition AV). Seven azimuthal source positions were reproduced, with loudspeakers visible at four azimuthal positions.

Results: In condition AV, the auditory images were perceived closer to the listener in rooms 2 and 3 than in room 1, with a larger effect in the reverberant than in the dry environment. In room 2, the perceived distance of the virtual sounds was more accurate in condition V than in conditions A and AV, where it was reduced. In room 3, differences in distance judgments between A, V, and AV conditions were much less pronounced. In contrast to distance, localization and compactness judgments were largely room independent, although localization judgments were less accurate and compactness ratings less consistent in conditions V and A than in condition VA. Conclusion: A mismatch between recording and playback room was found to be detrimental to virtual sound externalization. The auditory modality governed externalization in terms of perceived distance when cues from the recording and playback room were incongruent, whereby the auditory impression of the room was more critical the more reverberant the listening environment was. While the visual impression of the playback room did not affect perceived distance, visual cues helped resolve localization ambiguities and improved compactness perception.
The role of spectral detail in the binaural transfer function on perceived externalization in a reverberant environment

Individual binaural room impulse responses (BRIRs) were recorded at a distance of 1.5 m for azimuth angles of 0° and 50° in a reverberant room. Spectral details were reduced in either the direct or the reverberant part of the BRIRs by averaging the magnitude responses with band-pass filters. For various filter bandwidths, the modified BRIRs were convolved with broadband noise and listeners judged the perceived position of the noise when virtualized over headphones. Only reductions in spectral details of the direct part obtained with filter bandwidths broader than one equivalent rectangular bandwidth affected externalization. Reductions in spectral details of the reverberant part had only little influence on externalization. In both conditions, externalization was not as pronounced at 0° as at 50°. To characterize the auditory processes that may be involved in the perception of externalization, a quantitative model is proposed. The model includes an echo-suppression mechanism, a filterbank describing the frequency selectivity in the cochlea and a binaural stage that measures the deviations of the interaural level differences between the considered input and the unmodified input. These deviations, integrated across frequency, are then mapped to a value that corresponds to the perceived externalization.
Towards Objective Measures of Functional Hearing Abilities

Aims People with impaired hearing often have difficulties in hearing sounds in a noisy background. This problem is partially a result of the auditory systems reduced capacity to process temporal information in the sound signal. In this study we examined the relationships between perceptual sensitivity to temporal fine structure (TFS) cues, brainstem encoding of complex harmonic and amplitude modulated sounds, and the ability to understand speech in noise. Understanding these links will allow the development of an objective measure that could be used to detect changes in functional hearing before the onset of permanent threshold shifts. Methods We measured TFS sensitivity and speech in noise performance (QuickSIN) behaviourally in 34 normally hearing adults with ages ranging from 18 to 63 years. We recorded brainstem responses to complex harmonic sounds and a 4000 Hz carrier signal modulated at 110 Hz. We performed cross correlations between the stimulus waveforms and scalp-recorded brainstem responses to generate a simple measure of stimulus encoding accuracy, and correlated these measures with age, TFS sensitivity and speech-in-noise performance. Results Speech-in-noise performance was positively correlated with TFS sensitivity, and negatively correlated with age. TFS sensitivity was also positively correlated with stimulus encoding accuracy for the complex harmonic stimulus, while increasing age was associated with lower stimulus encoding accuracy for the modulated tone stimulus. Conclusions The results show that even in a group of people with normal hearing, increasing age was associated with reduced speech understanding, reduced TFS sensitivity, and reduced stimulus encoding accuracy (for the modulated tone stimulus). People with good TFS sensitivity also generally had less faithful brainstem encoding of a complex harmonic tone.
Validation of a Virtual Sound Environment System for Testing Hearing Aids

In the development process of modern hearing aids, test scenarios that reproduce natural acoustic scenes have become increasingly important in recent years for the evaluation of new signal processing algorithms. To achieve high ecological validity, such scenarios should include components like reverberation, background noise, and multiple interfering talkers. Loudspeaker-based sound field reproduction techniques, such as higher-order Ambisonics, allow for the simulation of such complex sound environments and can be used for realistic listening experiments with hearing aids. However, to successfully employ such systems, it is crucial to know how experimental results from a virtual environment translate to the corresponding real environment. In this study, speech reception thresholds (SRTs) were measured with normal-hearing listeners wearing hearing aids, both in a real room and in a simulation of that room auralized via a spherical array of 29 loudspeakers, using either Ambisonics or a nearest loudspeaker method. The benefit from a static beamforming algorithm was considered in comparison to a hearing aid setting with omnidirectional microphones. The measured SRTs were about 2-4 dB higher, and the benefit from the beamformer setting was, on average, about 1.5 dB smaller in the virtual room than in the real room. These differences resulted from a more diffuse sound field in the virtual room as indicated by differences in measured directivity patterns for the hearing aids and interaural cross-correlation coefficients. Overall, the considered VSE system may represent a valuable tool for testing the effects of hearing-aid signal processing on physical and behavioural outcome measures in realistic acoustic environments.
Variations in voice level and fundamental frequency with changing background noise level and talker-to-listener distance while wearing hearing protectors: A pilot study
Objective: Speech production in noise with varying talker-to-listener distance has been well studied for the open ear condition. However, occluding the ear canal can affect the auditory feedback and cause deviations from the models presented for the open-ear condition. Communication is a main concern for people wearing hearing protection devices (HPD). Although practical, radio communication is cumbersome, as it does not distinguish designated receivers. A smarter radio communication protocol must be developed to alleviate this problem. Thus, it is necessary to model speech production in noise while wearing HPDs. Such a model opens the door to radio communication systems that distinguish receivers and offer more efficient communication between persons wearing HPDs. Design: This paper presents the results of a pilot study aimed to investigate the effects of occluding the ear on changes in voice level and fundamental frequency in noise and with varying talker-to-listener distance. Study sample: Twelve participants with a mean age of 28 participated in this study. Results: Compared to existing data, results show a trend similar to the open ear condition with the exception of the occluded quiet condition. Conclusions: This implies that a model can be developed to better understand speech production for the occluded ear.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Université du Quebec, Institut National de la Recherche Scientifique
Authors: Bouserhal, R. E. (Ekstern), MacDonald, E. (Intern), Falk, T. H. (Ekstern), Voix, J. (Ekstern)
Pages: S13–S20
Publication date: 2016
Main Research Area: Technical/natural sciences

Publication information
Journal: International Journal of Audiology
Volume: 55
ISSN (Print): 1499-2027
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 2.07 SJR 1.289 SNIP 1.245
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 1.191 SNIP 1.217 CiteScore 1.79
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 1.3 SNIP 1.273 CiteScore 1.89
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 1.191 SNIP 1.499 CiteScore 1.94
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 1.232 SNIP 1.296 CiteScore 1.79
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 1.29 SNIP 1.209 CiteScore 1.78
ISI indexed (2011): ISI indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 1.075 SNIP 1.118
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 1.179 SNIP 1.11
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 1
Scopus rating (2008): SJR 1.217 SNIP 1.372
What affects envelope coding in the electrically stimulated auditory nerve?

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Joshi, S. N. (Intern), Dau, T. (Intern), Epp, B. (Intern)
Number of pages: 1
Publication date: 2016
Event: Poster session presented at 39th midwinter meeting of Association of Research in Otolaryngology, San Diego, CA, United States.
Main Research Area: Technical/natural sciences
Electronic versions:
2016_ARO_Joshi_et_al_1_.pdf
Publication: Research - peer-review › Poster – Annual report year: 2016

A machine-hearing system exploiting head movements for binaural sound localisation in reverberant conditions
This paper is concerned with machine localisation of multiple active speech sources in reverberant environments using two (binaural) microphones. Such conditions typically present a problem for ‘classical’ binaural models. Inspired by the human ability to utilise head movements, the current study investigated the influence of different head movement strategies on binaural sound localisation. A machine-hearing system that exploits a multi-step head rotation strategy for sound localisation was found to produce the best performance in simulated reverberant acoustic space. This paper also reports the public release of a free binaural room impulse responses (BRIRs) database that allows the simulation of head rotation used in this study.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Sheffield, Technische Universität Berlin
Authors: May, T. (Intern), Ma, N. (Ekstern), Wierstorf, H. (Ekstern), Brown, G. (Ekstern)
Pages: 2699-2703
Publication date: 2015

Host publication information
Title of host publication: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2015)
Publisher: IEEE
ISBN (Print): 9781467369978
Main Research Area: Technical/natural sciences
Conference: 40th IEEE International Conference on Acoustics, Speech and Signal Processing, Brisbane, Australia, 19/04/2015 - 19/04/2015
Head movements, Binaural localisation, Machine hearing, Reverberation, Binaural room impulse response
Electronic versions:
MaMayWierstorfBrown_ICASSP2015.pdf
DOIs:
10.1109/ICASSP.2015.7178461
A model of auditory nerve responses to electrical stimulation

Cochlear implants (CI) stimulate the auditory nerve (AN) with a train of symmetric biphasic current pulses comprising of a cathodic and an anodic phase. The cathodic phase is intended to depolarize the membrane of the neuron and to initiate an action potential (AP) and the anodic phase to neutralize the charge induced during the cathodic phase. Single-neuron recordings in cat auditory nerve using monophasic electrical stimulation show, however, that both phases in isolation can generate an AP. The site of AP generation differs for both phases, being more central for the anodic phase and more peripheral for the cathodic phase. This results in an average difference of 200 μs in spike latency for AP generated by anodic vs cathodic pulses. It is hypothesized here that this difference is large enough to corrupt the temporal coding in the AN. To quantify effects of pulse polarity on auditory perception of CI listeners, a model needs to incorporate the correct responsiveness of the AN to anodic and cathodic polarity. Previous models of electrical stimulation have been developed based on AN responses to symmetric biphasic stimulation or to monophasic cathodic stimulation. These models, however, fail to correctly predict responses to anodic stimulation. This study presents a model that simulates AN responses to anodic and cathodic stimulation. The main goal was to account for the data obtained with monophasic electrical stimulation in cat AN. The model is based on an exponential integrate-and-fire neuron with two partitions responding individually to anodic and cathodic stimulation. Membrane noise was parameterized based on reported relative spread of AN neurons. Firing efficiency curves and spike-latency distributions were simulated for monophasic and symmetric biphasic stimulation. The simulations were in line with the average data for firing thresholds and spike latencies for both, monophasic anodic and monophasic cathodic stimulation. The model also correctly predicted the shift in latency as a function of stimulation level. With the ability to account for the responsiveness to cathodic and anodic phases of electrical stimulation, this model can be applied to account for the response to arbitrary pulse shapes. The evaluation of the neural response to symmetric biphasic pulses helps to estimate the mutual interaction between the two pulse phases. A successful model can be generalized as a framework to test various stimulation strategies and to quantify their effect on the performance of CI listeners in psychophysical tasks.
considered. Here, a sinusoidally modulated WTN model accounting for temporal AM variations was used to generate realistic artificial stimuli in which the AM depth, frequency, and type, while determined from real on-site recordings, could be varied systematically. Subjective listening tests with such stimuli showed that a reduction in AM depth, quantified by the modulation depth spectrum, led to a significant decrease in annoyance. When the spectrotemporal characteristics of the original far-field stimuli were included in the model and the temporal AM variations were taken into account by varying the modulation index over time, neither AM frequency nor AM type were found to significantly affect annoyance. These findings suggest that the effect of AM parameters on WTN annoyance may depend on the intermittent nature of WTN AM

General information
State: Published
Organisations: Department of Electrical Engineering, Acoustic Technology, Hearing Systems
Authors: Ioannidou, C. (Intern), Santurette, S. (Intern), Jeong, C. (Intern)
Publication date: 2015
Main Research Area: Technical/natural sciences
Spectral properties
DOIs:
10.1121/1.4920611
Source: PublicationPreSubmission
Source-ID: 108627598
Publication: Research - peer-review › Paper – Annual report year: 2015

Apparent source width perception in normal-hearing, hearing-Impaired and aided listeners

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Widex A/S
Authors: Käsbach, J. (Intern), Wiinberg, A. (Intern), May, T. (Intern), Jepsen, M. L. (Ekstern), Dau, T. (Intern)
Number of pages: 4
Publication date: 2015
Host publication information
Title of host publication: Forschritte der Akustik DAGA'15
Main Research Area: Technical/natural sciences
Conference: DAGA 2015, Nürnberg, Germany, 16/03/2015 - 16/03/2015
Source: PublicationPreSubmission
Source-ID: 112764827
Publication: Research - peer-review › Article in proceedings – Annual report year: 2015

Auditory correlates of stimulus-induced variability in consonant perception

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Zaar, J. (Intern), Dau, T. (Intern)
Pages: 1023-1026
Publication date: 2015
Host publication information
Title of host publication: Forschritte der Akustik DAGA’15
Main Research Area: Technical/natural sciences
Conference: DAGA 2015, Nürnberg, Germany, 16/03/2015 - 16/03/2015
Source: PublicationPreSubmission
Source-ID: 110604466
Publication: Research - peer-review › Article in proceedings – Annual report year: 2015

Auditory Processing of Temporal Fine Structure: Effects of Age and Hearing Loss: Book Review

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, The Bionics Institute
Authors: Innes-Brown, H. (Ekstern), Marozeau, J. (Intern)
Number of pages: 1
Pages: 525
Can auditory study-state responses reflect place-specific cochlear dispersion?

The cochlear travelling wave propagates from the base to the apex, resulting in an increasing phase with distance from the cochlear base. Together with the tonotopic organization of the cochlea, this results in a frequency dependent delay of the resonance, a phenomenon known as cochlear dispersion. Previous studies showed the applicability of auditory evoked potentials (AEP) to investigate cochlear dispersion along the basilar membrane (BM) (e.g. Dau et al., 2000). In contrast to those studies, the present study maximizes the response in a given frequency region, aiming to objectively estimate local cochlear dispersion in humans.

Comparison of binaural microphones for externalization of sounds

Ubiquitous availability of media content through portable devices like media players and smartphones has resulted in an immensely increased popularity of headphones in recent years. However, while conventional stereo recordings usually create a good sense of space when listened to through loudspeakers, the sounds tend to be perceived inside the head (internalized) when headphones are used for listening. A more natural perception in headphone listening with sounds being perceived outside the head (externalized) can be achieved when recordings are made with dummy head microphones or with microphones placed inside the ear canals of a person. In this study, binaural room impulse responses (BRIRs) were measured with several commercially available binaural microphones, both placed inside the listeners' ears (individual BRIR) and on a head and torso simulator (generic BRIR). The degree of externalization of speech and noise stimuli was tested in a listening experiment with a multi-stimulus test. No influence was found for the stimulus signal, but the externalization scores were found to be lower for 0° incidence. With all microphones, relatively high externalization scores were achieved, and for all but one microphone, individual BRIRs resulted in slightly better externalization than generic ones.
Compensation of F0 and formant frequencies in a real-time pitch-perturbation paradigm

While producing speech, talkers monitor both somatosensory and auditory feedback. Many studies have demonstrated that if auditory feedback is manipulated in real-time (e.g., using an effects processor to shift the frequency spectrum), subjects compensate by modifying their F0 in the direction opposite to the perturbation. However, shifting the entire frequency spectrum alters both F0 and formant frequencies. While compensations for real-time formant perturbations have been previously observed, these studies have used a paradigm that is very different from that of traditional pitch-perturbation experiments. In the present study, compensations in both F0 and formant frequencies were compared for perturbations of sustained vowels using a traditional pitch-perturbation paradigm. Within a sustained utterance, the auditory feedback was shifted by a constant magnitude for a short duration. Previous studies have suggested that the large variability in compensation across individuals may be due to individual differences in weighting somatosensory and auditory feedback. Following this hypothesis, individuals’ compensations in F0 and formant frequency should be correlated. Results from the present experiment are discussed in this context and formant compensations are compared to results from experiments using a traditional formant-perturbation paradigm.

Cortical pitch representations of complex tones in musicians and non-musicians

Musicians typically show enhanced pitch-discrimination ability compared to non-musicians, consistent with the fact that musicians are more sensitive to some acoustic features critical for both speech and music processing. However, it is still unclear which mechanisms underlie this perceptual enhancement. In a previous behavioral study, musicians showed an increased pitch-discrimination performance for both resolved and unresolved complex tones suggesting an enhanced neural representation of pitch at central stages of the auditory system. The aim of this study was to clarify whether musicians show (i) differential neural activation in response to complex tones as compared to non-musicians and/or (ii) finer fundamental frequency (F0) representation in the auditory cortex. Assuming that the right auditory cortex is specialized in processing fine spectral changes, we hypothesized that an enhanced F0 representation in musicians would be associated with a stronger right-lateralized response to complex tones compared to non-musicians. Fundamental frequency (F0) discrimination thresholds were obtained for harmonic complex tones with F0s of 100 and 500 Hz, filtered in either a low or a high frequency region to vary the resolvability of audible harmonics. A sparse-sampling eventrelated functional magnetic resonance imaging (fMRI) paradigm was used to measure neural activation in all listeners while performing the same pitch-discrimination task for conditions of varying resolvability. The task difficulty was individually adjusted according to the previously obtained F0 discrimination thresholds. Preliminary results from 6 listeners (3 musicians and 3 non-musicians) showed that the behavioral discrimination thresholds of musicians were, on average, lower than the thresholds of non-musicians by about a factor of 2.3, independent of harmonic resolvability. A group analysis on the 6 listeners revealed no differential neural activation for resolved vs unresolved conditions, suggesting that cortical responses did not increase with increasing stimulus resolvability, when adjusting for the task difficulty across conditions and participants. A significant effect of processing demand, i.e., task demand estimated from both stimulus resolvability and task difficulty, was observed in both auditory cortices, with a larger neural activation in the right auditory region. Additionally, no differential activation was observed in the musicians vs. the non-musicians. Overall, these preliminary findings suggest an involvement of a postero-lateral region in both auditory cortices during a pitch-discrimination task with conditions of varying processing demand. Cortical responses were larger in the right than in the left auditory cortex, suggesting an increasing activation of the right-lateralized pitch-sensitive cortical areas with increasing taskprocessing demand.
Cortical pitch representations of complex tones in musicians and non-musicians

Musicians have been shown to have an enhanced pitch-discrimination ability compared to non-musicians for complex tones with either resolved or unresolved harmonics [1, 2, 3, 4, 5]. It is unclear whether this perceptual enhancement can be ascribed to an enhanced neural representation of pitch at central stages of the auditory system. The aim of this study was to clarify whether (i) cortical responses increase with harmonic resolvability, as suggested in previous studies [6, 7], and whether musicians show (ii) differential neural activation in response to complex tones as compared to non-musicians and/or (iii) a finer fundamental frequency (F0) representation in the auditory cortex. Assuming that the right auditory cortex is specialized in processing fine spectral changes, we hypothesized that an enhanced F0 representation in musicians would be associated with a stronger right-lateralized response to complex tones compared to non-musicians.

Dichotic Listening Can Improve Perceived Clarity of Music in Cochlear Implant Users

Musical enjoyment for cochlear implant (CI) recipients is often reported to be unsatisfactory. Our goal was to determine whether the musical experience of postlingually deafened adult CI recipients could be enriched by presenting the bass and treble clef parts of short polyphonic piano pieces separately to each ear (dichotic). Dichotic presentation should artificially enhance the lateralization cues of each part and help the listeners to better segregate them and thus provide greater clarity. We also hypothesized that perception of the intended emotion of the pieces and their overall enjoyment would be enhanced in the dichotic mode compared with the monophonic (both parts in the same ear) and the diotic mode (both parts in both ears). Twenty-eight piano pieces specifically composed to induce sad or happy emotions were selected. The tempo of the pieces, which ranged from lento to presto covaried with the intended emotion (from sad to happy). Thirty participants (11 normal-hearing listeners, 11 bimodal CI and hearing-aid users, and 8 bilaterally implanted CI users) participated in this study. Participants were asked to rate the perceived clarity, the intended emotion, and their preference of each piece in different listening modes. Results indicated that dichotic presentation produced small significant improvements in subjective ratings based on perceived clarity. We also found that preference and clarity ratings were significantly higher for pieces with fast tempi compared with slow tempi. However, no significant differences between diotic and dichotic presentation were found for the participants’ preference ratings, or their judgments of intended emotion.
Effect of harmonic rank on the streaming of complex tones
The effect of the rank of the harmonics on sequential stream segregation of complex tones was investigated for normal-hearing participants with no musical training. It was hypothesized that stream segregation would be greater for tones with high pitch salience, as assessed by fundamental frequency (f0) difference limens. Pitch salience is highest for tones containing some low (resolved) harmonics, but is also fairly high for tones containing harmonics of intermediate rank. The tones were bandpass filtered between 2 and 4 kHz and harmonic rank was varied by changing the f0. There was a significant trend for less stream segregation with increasing harmonic rank. The amount of stream segregation was inversely correlated with the f0 difference limens, consistent with the hypothesis.
Effects of cochlear compression and frequency selectivity on pitch discrimination of complex tones with unresolved harmonics

Physiological studies have shown that noise-induced sensorineural hearing loss (SNHL) enhances the amplitude of envelope coding in auditory-nerve fibers. As pitch coding of unresolved complex tones is assumed to rely on temporal envelope coding mechanisms, this study investigated pitchdiscrimination performance in listeners with SNHL. Pitch-discrimination thresholds were obtained in 14 normal-hearing (NH) and 10 hearingimpaired (HI) listeners for sine-phase (SP) and random-phase (RP) unresolved complex tones. The HI listeners performed, on average, similarly as the NH listeners in the SP condition and worse than NH listeners in the RP condition. Cochlear compression and auditory filter bandwidths were estimated in the same listeners. A significant correlation was found between the reduction of cochlear compression and the difference between RP and SP pitch-discrimination thresholds. The effects of degraded frequency selectivity and loss of compression were considered in a model as potential factors in envelope enhancement. The model revealed that a broadening of the auditory filters led to an increase of the modulation depth in the SP condition, while it did not have any effect for the RP condition. Overall, these findings suggest that both reduced cochlear compression and auditory filter broadening alter the envelope representation of unresolved complex tones, leading to changes in pitch-discrimination performance.

Effects of harmonic roving on pitch discrimination

Performance in pitch discrimination tasks is limited by variability intrinsic to listeners which may arise from peripheral auditory coding limitations or more central noise sources. Perceptual limitations may be characterized by measuring an observer’s change in performance when introducing external noise in the physical stimulus (Lu and Dosher, 2008). The present study used this approach to attempt to quantify the “internal noise” involved in pitch coding of harmonic complex tones by estimating the amount of harmonic roving required to impair pitch discrimination performance. It remains a matter of debate whether pitch perception of natural complex sounds mostly relies on either spectral excitation-based information or temporal periodicity information. Comparing the way internal noise affects the internal representations of such information to how it affects pitch discrimination performance may help clarify pitch coding mechanisms. As training on frequency discrimination tasks has been found to result in a reduction of internal noise (Jones et al., 2013), it was also investigated whether the effect of harmonic roving varied with musical training.
Effects of harmonic roving on pitch discrimination

Performance in pitch discrimination tasks is limited by variability intrinsic to listeners which may arise from peripheral auditory coding limitations or more central noise sources. The present study aimed at quantifying such "internal noise" by estimating the amount of harmonic roving required to impair pitch discrimination performance. Fundamental-frequency difference limens (F0DLs) were obtained in normal-hearing listeners with and without musical training for complex tones filtered between 1.5 and 3.5 kHz with F0s of 300 Hz (resolved harmonics) and 75 Hz (unresolved harmonics). The harmonicity of the tone complexes was varied by systematically roving the frequency of individual harmonics, which was taken from a Gaussian distribution centered on the nominal frequency in every stimulus presentation. The amount of roving was determined by the standard deviation of this distribution, which varied between 0% and 16% of the tested F0. F0DLs for resolved harmonics remained unaffected for up to 6% roving, and increased thereafter. For unresolved harmonics, performance remained stable up to larger roving values. The results demonstrate a systematic relationship between F0DLs and stimulus variability that could be used to quantify the internal noise and provide strong constraints for physiologically inspired models of pitch perception.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Technical University of Denmark
Authors: Santurette, S. (Intern), de Kérangal, M. L. G. (Ekstern), Joshi, S. N. (Intern)
Pages: 2226
Publication date: 2015
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of the Acoustical Society of America
Volume: 137
Issue number: 4
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Web of Science (2009): Indexed yes
Effects of incongruent auditory and visual room-related cues on sound externalization

Sounds presented via headphones are typically perceived inside the head. However, the illusion of a sound source located out in space away from the listener's head can be generated with binaural headphone-based auralization systems by convolving anechoic sound signals with a binaural room impulse response (BRIR) measured with miniature microphones placed in the listener's ear canals. Sound externalization of such virtual sounds can be very convincing and robust but there have been reports that the illusion might break down when the listening environment differs from the room in which the BRIRs were recorded [1,2,3]. This may be due to incongruent auditory cues between the recording and playback room during sound reproduction [2]. Alternatively, an expectation effect caused by the visual impression of the room may affect the position of the perceived auditory image [3]. Here, we systematically investigated whether incongruent auditory and visual room-related cues affected sound externalization in terms of perceived distance, azimuthal localization, and compactness.

General information

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Technical University of Denmark
Authors: Carvajal, J. C. G. (Ekstern), Santurette, S. (Intern), Cubick, J. (Intern), Dau, T. (Intern)
Publication date: 2015
Event: Poster session presented at Tenth anniversary symposium of the international laboratory for Brain, Music, and Sound Research, Montréal, Canada.
Main Research Area: Technical/natural sciences
Electronic versions:
Carvajal2015.pdf
Source: PublicationPreSubmission
Source-ID: 117493959
Publication: Research - peer-review › Poster – Annual report year: 2015

Effects of manipulating the signal-to-noise envelope power ratio on speech intelligibility

Jørgensen and Dau [(2011). J. Acoust. Soc. Am. 130, 1475–1487] suggested a metric for speech intelligibility prediction based on the signal-to-noise envelope power ratio (SNRenv), calculated at the output of a modulation-frequency selective process. In the framework of the speech-based envelope power spectrum model (sEPSM), the SNRenv was demonstrated to account for speech intelligibility data in various conditions with linearly and nonlinearly processed noisy speech, as well as for conditions with stationary and fluctuating interferers. Here, the relation between the SNRenv and speech intelligibility was investigated further by systematically varying the modulation power of either the speech or the noise before mixing the two components, while keeping the overall power ratio of the two components constant. A good correspondence between the data and the corresponding sEPSM predictions was obtained when the noise was
manipulated and mixed with the unprocessed speech, consistent with the hypothesis that SNRenv is indicative of speech intelligibility. However, discrepancies between data and predictions occurred for conditions where the speech was manipulated and the noise left untouched. In these conditions, distortions introduced by the applied modulation processing were detrimental for speech intelligibility, but not reflected in the SNRenv metric, thus representing a limitation of the modeling framework.

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Jørgensen, S. (Intern), Decorsière, R. J. B. (Intern), Dau, T. (Intern)
Pages: 1401–1410
Publication date: 2015
Main Research Area: Technical/natural sciences

**Publication information**
Journal: Journal of the Acoustical Society of America
Volume: 137
Issue number: 3
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 0.848 SNIP 1.633
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 0.865 SNIP 1.647
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 0.752 SNIP 1.559
Effects of musical training on pitch discrimination of resolved and unresolved complex tones

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Bianchi, F. (Intern), Santurette, S. (Intern), Wendt, D. (Intern), Dau, T. (Intern)
Number of pages: 1
Publication date: 2015
Event: Poster session presented at 38th Annual MidWinter Meeting of the Association for Research in Otolaryngology, Baltimore, MD, United States.
Main Research Area: Technical/natural sciences
Electronic versions:
Bianchi_et_al_2015_ARO.pdf
Source: PublicationPreSubmission
Source-ID: 111863276
Publication: Research - peer-review › Poster – Annual report year: 2015

Effects of syntactic complexity on word recognition in noise

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Copenhagen
Authors: Kristensen, L. B. (Ekstern), Wendt, D. (Intern)
Publication date: 2015
Event: Poster session presented at 7th Workshop on Speech in Noise, Copenhagen, Denmark.
Main Research Area: Technical/natural sciences
Source: FindIt
Source-ID: 2280083073
Publication: Research - peer-review › Poster – Annual report year: 2015

Evaluation of a clinical auditory profile in hearing-aid candidates

Hearing-impaired (HI) listeners often complain about communicating in the presence of background noise, although audibility may be restored by a hearing-aid (HA). The audiogram typically forms the basis for HA fitting, such that people with similar audiograms are given the same prescription by default. However, this does not necessarily lead to the same HA benefit. This study aimed at identifying clinically relevant tests that may be informative in addition to the audiogram and relate more directly to HA benefit. Twenty-nine HI listeners performed fast tests of loudness perception, spectral and temporal resolution, binaural hearing, speech intelligibility in stationary and fluctuating noise, and a working-memory test. Six weeks after HA fitting they answered the International Outcome Inventory – Hearing Aid evaluation. The HI group was homogeneous based on the audiogram, but only one test was correlated to pure-tone hearing thresholds. Moreover, HI listeners who took the least advantage from fluctuations in background noise in terms of speech intelligibility experienced...
gradual HA benefit. Further analysis of whether specific outcomes are directly related to speech intelligibility in fluctuating noise could be relevant for concrete HA fitting applications.

**Evaluation of peripheral compression and auditory nerve fiber intensity coding using auditory steady-state responses (ASSR)**

The compressive nonlinearity of the auditory system is assumed to be an epiphenomenon of a healthy cochlea and, particularly, of outer-hair cell function. Another ability of the healthy auditory system is to enable communication in acoustical environments with high-level background noises. Evaluation of these properties provides information about the health state of the system. It has been shown that a loss of outer hair cells leads to a reduction in peripheral compression. It has also recently been shown in animal studies that noise over-exposure, producing temporary threshold shifts, can cause auditory nerve fiber (ANF) deafferentation in predominantly low-spontaneous rate (SR) fibers. In the present study, auditory steadystate response (ASSR) level growth functions were measured to evaluate the applicability of ASSR to assess compression and the ability to code intensity fluctuations at high stimulus levels. Level growth functions were measured in normal-hearing adults at stimulus levels ranging from 20 to 90 dB SPL. To evaluate compression, ASSR were measured for multiple carrier frequencies simultaneously. To evaluate intensity coding at high intensities, ASSR were measured using a single carrier frequency at four modulation depths between 25 and 100%. The data showed that ASSR level growth functions exhibited compression of about 0.25 dB/dB. For levels above 60 dB SPL, the slope showed higher variability for the different modulation depths across subjects than for lower levels. The results indicate that the slope of the ASSR level growth function can be used to estimate peripheral compression simultaneously at four frequencies below 60 dB SPL, while the slope above 60 dB SPL may provide information about the integrity of intensity coding of low-SR fibers.
Exploiting deep neural networks and head movements for binaural localisation of multiple speakers in reverberant conditions

This paper presents a novel machine-hearing system that exploits deep neural networks (DNNs) and head movements for binaural localisation of multiple speakers in reverberant conditions. DNNs are used to map binaural features, consisting of the complete crosscorrelation function (CCF) and interaural level differences (ILDs), to the source azimuth. Our approach was evaluated using a localisation task in which sources were located in a full 360-degree azimuth range. As a result, front-back confusions often occurred due to the similarity of binaural features in the front and rear hemifields. To address this, a head movement strategy was incorporated in the DNN-based model to help reduce the front-back errors. Our experiments show that, compared to a system based on a Gaussian mixture model (GMM) classifier, the proposed DNN system substantially reduces localisation errors under challenging acoustic scenarios in which multiple speakers and room reverberation are present.

Extraction of OAEs During Multi-Frequency ASSR Recordings With the Goal to Estimate Peripheral Compression

Auditory steady-state responses (ASSR) allow objective assessment of auditory function using electroencephalography (EEG). If peripheral compression is, at least partially, due to cochlear compression, the amplitude of ASSR as a function of level can be assumed to reflect the compressive growth of the cochlear nonlinearity. A recent study (Encina Llamas et al., ARO2014) showed, that compressive inputoutput functions with slopes similar to proposed compression ratios of cochlear level-growth functions can be found using ASSR obtained by stimulation with multiple sinusoidally-amplitude-modulated (SAM) tones. The goal of the current study is to investigate the estimation of cochlear compression using the same stimuli as used for ASSR (SAM tones).
Face configuration affects speech perception: Evidence from a McGurk mismatch negativity study

We perceive identity, expression and speech from faces. While perception of identity and expression depends crucially on the configuration of facial features it is less clear whether this holds for visual speech perception.

Facial configuration is poorly perceived for upside-down faces as demonstrated by the Thatcher illusion in which the orientation of the eyes and mouth with respect to the face is inverted (Thatcherization). This gives the face a grotesque appearance but this is only seen when the face is upright.

Thatcherization can likewise disrupt visual speech perception but only when the face is upright indicating that facial configuration can be important for visual speech perception. This effect can propagate to auditory speech perception through audiovisual integration so that Thatcherization disrupts the McGurk illusion in which visual speech perception alters perception of an incongruent acoustic phoneme. This is known as the McThatcher effect.

Here we show that the McThatcher effect is reflected in the McGurk mismatch negativity (MMN). The MMN is an event-related potential elicited by a change in auditory perception. The McGurk-MMN can be elicited by a change in auditory perception due to the McGurk illusion without any change in the acoustic stimulus.

We found that Thatcherization disrupted a strong McGurk illusion and a correspondingly strong McGurk-MMN only for upright faces. This confirms that facial configuration can be important for audiovisual speech perception. For inverted faces we found a weaker McGurk illusion but, surprisingly, no MMN. We also found no correlation between the strength of the McGurk illusion and the amplitude of the McGurk-MMN. We suggest that this may be due to a threshold effect so that a strong McGurk illusion is required to elicit the McGurk-MMN.

General information

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems, Hearing Systems, Department of Electrical Engineering
Authors: Eskelund, K. (Intern), MacDonald, E. (Intern), Andersen, T. (Intern)
Pages: 48-54
Publication date: 2015
Main Research Area: Technical/natural sciences

Publication information

Journal: Neuropsychologia
Volume: 66
ISSN (Print): 0028-3932
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Scopus rating (2016): SJR 1.938 SNIP 1.12 CiteScore 3.34
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 2.054 SNIP 1.152 CiteScore 3.42
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 2.289 SNIP 1.31 CiteScore 3.82
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 2.427 SNIP 1.356 CiteScore 4.09
ISI indexed (2013): ISI indexed yes
Fast assessment of auditory spectral and temporal resolution

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Technical University of Denmark
Authors: Pelzer, A. (Ekstern), Santurette, S. (Intern), Bianchi, F. (Intern), Dau, T. (Intern)
Pages: 932-935
Publication date: 2015

Host publication information
Title of host publication: Forschritte der Akustik DAGA’15
Main Research Area: Technical/natural sciences
Conference: DAGA 2015, Nürnberg, Germany, 16/03/2015 - 16/03/2015
Source: PublicationPreSubmission
Source-ID: 112762328
Publication: Research - peer-review › Article in proceedings – Annual report year: 2015

Formant compensation for auditory feedback with English vowels
Past studies have shown that speakers spontaneously adjust their speech acoustics in response to their auditory feedback perturbed in real time. In the case of formant perturbation, the majority of studies have examined speaker's compensatory production using the English vowel /ɛ/ as in the word “head.” Consistent behavioral observations have been reported, and there is lively discussion as to how the production system integrates auditory versus somatosensory feedback to control vowel production. However, different vowels have different oral sensation and proprioceptive information due to differences in the degree of lingual contact or jaw openness. This may in turn influence the ways in which speakers compensate for auditory feedback. The aim of the current study was to examine speakers' compensatory behavior with six English monophthongs. Specifically, the current study tested to see if "closed vowels" would show less compensatory production than "open vowels" because closed vowels' strong lingual sensation may richly specify production via somatosensory feedback. Results showed that, indeed, speakers exhibited less compensatory production with the closed
vowels. Thus sensorimotor control of vowels is not fixed across all vowels; instead it exerts different influences across different vowels.

**General information**

State: Published

Organisations: Department of Electrical Engineering, Hearing Systems, Western University, Queen's University

Authors: Mitsuya, T. (Ekstern), MacDonald, E. N. (Intern), Munhall, K. G. (Ekstern), Purcell, D. W. (Ekstern)

Pages: 413-424

Publication date: 2015

Main Research Area: Technical/natural sciences

**Publication information**

Journal: Journal of the Acoustical Society of America

Volume: 138

Issue number: 1

ISSN (Print): 0001-4966

Ratings:

BFI (2018): BFI-level 2

Web of Science (2018): Indexed yes

BFI (2017): BFI-level 2

Web of Science (2017): Indexed yes

BFI (2016): BFI-level 2

Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27

Web of Science (2016): Indexed yes

BFI (2015): BFI-level 2

Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77

Web of Science (2015): Indexed yes

BFI (2014): BFI-level 2

Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8

Web of Science (2014): Indexed yes

BFI (2013): BFI-level 2

Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2

ISI indexed (2013): ISI indexed yes

Web of Science (2013): Indexed yes

BFI (2012): BFI-level 2

Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75

ISI indexed (2012): ISI indexed yes

Web of Science (2012): Indexed yes

BFI (2011): BFI-level 2

Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68

ISI indexed (2011): ISI indexed yes

Web of Science (2011): Indexed yes

BFI (2010): BFI-level 2

Scopus rating (2010): SJR 0.754 SNIP 1.528

Web of Science (2010): Indexed yes

BFI (2009): BFI-level 2

Scopus rating (2009): SJR 0.783 SNIP 1.717

Web of Science (2009): Indexed yes

BFI (2008): BFI-level 2

Scopus rating (2008): SJR 0.848 SNIP 1.633

Web of Science (2008): Indexed yes

Scopus rating (2007): SJR 0.865 SNIP 1.647

Web of Science (2007): Indexed yes

Scopus rating (2006): SJR 0.752 SNIP 1.559

Web of Science (2006): Indexed yes

Scopus rating (2005): SJR 0.954 SNIP 1.749
How Hearing Impairment Affects Sentence Comprehension: Using Eye Fixations to Investigate the Duration of Speech Processing

The main objective of this study was to investigate the extent to which hearing impairment influences the duration of sentence processing. An eye-tracking paradigm is introduced that provides an online measure of how hearing impairment prolongs processing of linguistically complex sentences; this measure uses eye fixations recorded while the participant listens to a sentence. Eye fixations toward a target picture (which matches the aurally presented sentence) were measured in the presence of a competitor picture. Based on the recorded eye fixations, the single target detection amplitude, which reflects the tendency of the participant to fixate the target picture, was used as a metric to estimate the duration of sentence processing. The single target detection amplitude was calculated for sentence structures with different levels of linguistic complexity and for different listening conditions: in quiet and in two different noise conditions. Participants with hearing impairment spent more time processing sentences, even at high levels of speech intelligibility. In addition, the relationship between the proposed online measure and listener-specific factors, such as hearing aid use and cognitive abilities, was investigated. Longer processing durations were measured for participants with hearing impairment who were not accustomed to using a hearing aid. Moreover, significant correlations were found between sentence processing duration and individual cognitive abilities (such as working memory capacity or susceptibility to interference). These findings are discussed with respect to audiological applications.

General information

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Oldenburg
Authors: Wendt, D. (Intern), Kollmeier, B. (Ekstern), Brand, T. (Ekstern)
Pages: 1-18
Publication date: 2015
Main Research Area: Technical/natural sciences

Publication information

Journal: Trends in Hearing
Volume: 19
ISSN (Print): 2331-2165
Ratings:
Web of Science (2018): Indexed yes
Web of Science (2017): Indexed Yes
Scopus rating (2016): CiteScore 3.61
Web of Science (2016): Indexed yes
Scopus rating (2015): SJR 1.608 SNIP 1.468 CiteScore 2.21
Web of Science (2015): Indexed yes
Scopus rating (2014): SJR 1.006 SNIP 1.099 CiteScore 0
Web of Science (2014): Indexed yes
Scopus rating (2013): SJR 0.953 SNIP 1.009
Web of Science (2012): Indexed yes
Scopus rating (2012): SJR 0.981 SNIP 0.909
Impact of background noise and sentence complexity on cognitive processing demands

Speech comprehension in adverse listening conditions requires cognitive processing demands. Processing demands can increase with acoustically degraded speech but also depend on linguistic aspects of the speech signal, such as syntactic complexity. In the present study, pupil dilations were recorded in 19 normal-hearing participants while processing sentences that were either syntactically simple or complex and presented in either high- or low-level background noise. Furthermore, the participants were asked to rate the subjectively perceived difficulty of sentence comprehension. The results showed that increasing noise levels had a greater impact on the perceived difficulty than sentence complexity. In contrast, the processing of complex sentences resulted in greater and more prolonged pupil dilations. The results suggest that while pupil dilations may correlate with cognitive processing demands, acoustic noise has a greater impact on the subjective perception of difficulty.

General information

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Wendt, D. (Intern), Dau, T. (Intern), Hjortkjær, J. (Intern)
Number of pages: 8
Publication date: 2015

Host publication information

Title of host publication: Proceedings of ISAAR 2015: Individual Hearing Loss – Characterization, Modelling
Editors: Santurette, S., Dau, T., Dalsgaard, J. C., Tranebjaerg, L., Andersen, T.
ISBN (Print): 978-87-990013-5-4
Main Research Area: Technical/natural sciences
Conference: 5th International Symposium on Auditory and Audiological Research, Nyborg, Denmark, 26/08/2015 - 26/08/2015
Electronic versions: 47_93_1_SM.pdf
Publication: Research - peer-review › Article in proceedings – Annual report year: 2016

Improvements in Elimination of Loudspeaker Distortion in Acoustic Measurements

This paper investigates the influence of nonlinear components that contaminate the linear response of acoustic transducers, and presents improved methods for eliminating the influence of nonlinearities in acoustic measurements. The method is evaluated with pure sinusoidal signals as well as swept sine signal and is tested on models of memoryless nonlinear systems as well as nonlinear loudspeakers. The method is shown to give a clear benefit over existing methods. Two techniques that improve the signal-to-noise ratio are demonstrated: the first uses more measurement levels than the number of orders to be separated, whereas the other one is based on standard Tikhonov regularization. Both methods are shown to significantly improve the signal-to-noise ratio.
General information
State: Published
Organisations: Department of Electrical Engineering, Acoustic Technology, Hearing Systems
Authors: Agerkvist, F. T. (Intern), Torras Rosell, A. (Intern), McWalter, R. I. (Intern)
Number of pages: 10
Publication date: 2015

Host publication information
Title of host publication: Proceedings of 138th International Audio Engineering Society (AES) Convention
Publisher: Audio Engineering Society
Main Research Area: Technical/natural sciences
Conference: 138th International Audio Engineering Society (AES) Convention, Warsaw, Poland, 07/05/2015 - 07/05/2015
Source: PublicationPreSubmission
Source-ID: 110964245
Publication: Research - peer-review › Article in proceedings – Annual report year: 2015

Informational interference from a competing talker: a thought-provoking but elusive construct
A competing talker can impair speech processing through both energetic masking and informational, cognitive aspects of masking. We refer to the latter as informational interference. We hypothesized that informational interference depletes processing resources that could otherwise be allocated to recognizing and understanding target speech. Consequently, informational interference should be more pronounced for target sentences with high processing demands (complex syntax) than for sentences with low processing demands (simple syntax). Furthermore, informational interference should be particularly marked when participants' own processing demands are increased, as with non-native listeners. Using a speeded picture selection task, we assessed native and non-native listeners' understanding of subject-relative (simple) and object-relative (complex) sentences, played against a competing talker vs. a matched energetic mask, at various signal-to-noise ratios (SNRs). Although object-relative sentences were more demanding than subject-relative sentences, the competing talker did not affect performance more than did energetic mask controls. This pattern was comparable for native and non-native listeners, and across SNRs. Moreover, individual differences in working memory were not related to differences in the speeded-selection task, regardless of the mask. Eye-tracking and pupillometric versions of this experiment also yielded similar results. Thus, contrary to prior research, we found no evidence that a competing talker requires greater processing resources than energetic masking alone. To address this discrepancy, an ongoing study aims to determine whether the semantic content of the competing talker's utterances modulates attention to the target.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Universidad del Pais Vasco, University of York
Authors: Valdés-Laribi, H. (Ekstern), Wendt, D. (Intern), MacDonald, E. (Intern), Cooke, M. (Ekstern), Mattys, S. (Ekstern)
Number of pages: 1
Publication date: 2015
Main Research Area: Technical/natural sciences
Electronic versions:
Informational_interference_from_a_competing_talker.pdf

Bibliographical note
Poster no. 83
Source: PublicationPreSubmission
Source-ID: 112335768
Publication: Research - peer-review › Poster – Annual report year: 2015

Inversion of Auditory Spectrograms, Traditional Spectrograms, and Other Envelope Representations
Envelope representations such as the auditory or traditional spectrogram can be defined by the set of envelopes from the outputs of a filterbank. Common envelope extraction methods discard information regarding the fast fluctuations, or phase, of the signal. Thus, it is difficult to invert, or reconstruct a time-domain signal from, an arbitrary envelope representation. To address this problem, a general optimization approach in the time domain is proposed here, which iteratively minimizes the distance between a target envelope representation and that of a reconstructed time-domain signal. Two implementations of this framework are presented for auditory spectrograms, where the filterbank is based on the behavior of the basilar membrane and envelope extraction is modeled on the response of inner hair cells. One implementation is direct while the other is a two-stage approach that is computationally simpler. While both can accurately invert an auditory spectrogram, the two-stage approach performs better on time-domain metrics. The same framework is applied to traditional spectrograms based on the magnitude of the short-time Fourier transform. Inspired by human perception of loudness, a modification to the framework is proposed, which leads to a more accurate inversion of traditional spectrograms.
Perceptual Interactions Between Electrodes Using Focused and Monopolar Cochlear Stimulation

In today's cochlear implant (CI) systems, the monopolar (MP) electrode configuration is the most commonly used stimulation mode, requiring only a single current source. However, with an implant that will allow simultaneous activation of multiple independent current sources, it is possible to implement an all-polar (AP) stimulation mode designed to create a focused electrical field. The goal of this experiment was to study the potential benefits of this all-polar mode for reducing uncontrolled electrode interactions compared with the monopolar mode. The five participants who took part in the study...
were implanted with a research device that was connected via a percutaneous connector to a benchtop stimulator providing 22 independent current sources. The perceptual effects of the AP mode were tested in three experiments. In Experiment 1, the current level difference between loudness-matched sequential and simultaneous stimuli composed of 2 spatially separated pulse trains was measured as function of the electrode separation. Results indicated a strong current-summation interaction for simultaneous stimuli in the MP mode for separations up to at least 4.8 mm. No significant interaction was found in the AP mode beyond a separation of 2.4 mm. In Experiment 2, a forward-masking paradigm was used with fixed equally loud probes in AP and MP modes, and AP maskers presented on different electrode positions. Results indicated a similar spatial masking pattern between modes. In Experiment 3, subjects were asked to discriminate between across-electrode temporal delays. It was hypothesized that discrimination would decrease with electrode separation faster in AP compared to MP modes. However, results showed no difference between the two modes. Overall, the results indicated that the AP mode produced less current spread than MP mode but did not lead to a significant advantage in terms of spread of neuronal excitation at equally loud levels.

**General information**

State: Published  
Organisations: Department of Electrical Engineering, Hearing Systems, University of Melbourne, Cochlear Ltd.  
Authors: Marozeau, J. (Intern), McDermott, H. J. (Ekstern), Swanson, B. A. (Ekstern), Mckay, C. M. (Ekstern)  
Pages: 401-412  
Publication date: 2015  
Main Research Area: Technical/natural sciences

**Publication information**

Journal: J A R O  
Volume: 16  
Issue number: 3  
ISSN (Print): 1525-3961  
Ratings:  
BFI (2018): BFI-level 2  
Web of Science (2018): Indexed yes  
BFI (2017): BFI-level 2  
Web of Science (2017): Indexed yes  
BFI (2016): BFI-level 2  
Scopus rating (2016): CiteScore 2.61 SJR 1.363 SNIP 1.269  
Web of Science (2016): Indexed yes  
BFI (2015): BFI-level 2  
Scopus rating (2015): SJR 1.688 SNIP 1.48 CiteScore 2.95  
Web of Science (2015): Indexed yes  
BFI (2014): BFI-level 2  
Scopus rating (2014): SJR 1.592 SNIP 1.371 CiteScore 2.84  
BFI (2013): BFI-level 2  
Scopus rating (2013): SJR 1.333 SNIP 1.432 CiteScore 2.67  
ISI indexed (2013): ISI indexed yes  
Web of Science (2013): Indexed yes  
BFI (2012): BFI-level 2  
Scopus rating (2012): SJR 1.535 SNIP 1.299 CiteScore 2.74  
ISI indexed (2012): ISI indexed yes  
BFI (2011): BFI-level 2  
Scopus rating (2011): SJR 1.416 SNIP 1.5 CiteScore 2.98  
ISI indexed (2011): ISI indexed yes  
BFI (2010): BFI-level 2  
Scopus rating (2010): SJR 1.617 SNIP 1.406  
Web of Science (2010): Indexed yes  
BFI (2009): BFI-level 2  
Scopus rating (2009): SJR 1.406 SNIP 1.17  
BFI (2008): BFI-level 2  
Scopus rating (2008): SJR 1.456 SNIP 1.066  
Scopus rating (2007): SJR 1.418 SNIP 1.123  
Scopus rating (2006): SJR 1.398 SNIP 1.074
Perceptual space induced by cochlear implant all-polar stimulation mode

It has often been argued that a main limitation of the cochlear implant is the spread of current induced by each electrode, which activates an inappropriately large range of sensory neurons. In order to reduce this spread, a new stimulation mode, the all-polar mode, was tested with 5 participants. It was designed to activate all the electrodes simultaneously with appropriate current levels and polarities to recruit narrower regions of auditory nerves in the region of specific intra-cochlear electrode positions (denoted all-polar electrodes). In this study, the all-polar mode was compared to the current commercial stimulation mode: the monopolar mode. The participants were asked to judge the sound dissimilarity between pairs of 2-electrode stimuli that differed in the electrode positions and were presented in either monopolar or all-polar mode. The dissimilarity ratings were analysed using a multidimensional scaling technique and a threedimensional stimulus perceptual space was produced. For both modes, the first perceptual dimension was highly correlated with the average position of the electrical stimulation and the second dimension moderately correlated with the distance between the two electrodes. The monopolar and all-polar stimuli were separated by a third dimension, which may indicate that allpolar stimuli have a perceptual quality that differs from monopolar stimuli.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, The Bionics Institute
Authors: Marozeau, J. (Intern), Mckay, C. M. (Ekstern)
Number of pages: 7
Pages: 333-339
Publication date: 2015

Host publication information
Title of host publication: Proceedings of ISAAR 2015
Editors: Santurette, S., Dau, T., Dalsgaard, J. C., Tranebjærg, L., Andersen, T.
ISBN (Print): 978-87-990013-5-4
Main Research Area: Technical/natural sciences
Conference: 5th International Symposium on Auditory and Audiological Research, Nyborg, Denmark, 26/08/2015 - 26/08/2015
Source: PublicationPreSubmission
Source-ID: 124208615
Publication: Research - peer-review › Article in proceedings – Annual report year: 2016

Relation between temporal envelope coding, pitch discrimination, and compression estimates in listeners with sensorineural hearing loss

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Bianchi, F. (Intern), Santurette, S. (Intern), Fereczkowski, M. (Intern), Dau, T. (Intern)
Number of pages: 1
Publication date: 2015
Event: Poster session presented at 169th Meeting of the Acoustical Society of America, Pittsburgh, Pa, United States.
Main Research Area: Technical/natural sciences
Relation between temporal envelope coding, pitch discrimination, and compression estimates in listeners with sensorineural hearing loss

Recent physiological studies in animals showed that noise-induced sensorineural hearing loss (SNHL) increased the amplitude of envelope coding in single auditory-nerve fibers. The present study investigated whether SNHL in human listeners was associated with enhanced temporal envelope coding, whether this enhancement affected pitch discrimination performance, and whether loss of compression following SNHL was a potential factor in envelope coding enhancement. Envelope processing was assessed in normal-hearing (NH) and hearing-impaired (HI) listeners in a behavioral amplitude-modulation detection task. Fundamental frequency difference limens (F0DLs) were obtained in the same listeners for complex tones with varying harmonic resolvability. Basilar-membrane input/output functions were measured to assess individual compression ratios. For NH listeners, F0DLs decreased with increasing harmonic resolvability. For the unresolved conditions, all five HI listeners performed as good as or better than NH listeners with matching musical experience. Two HI listeners showed lower amplitude-modulation detection thresholds than NH listeners for low modulation rates, and one of these listeners also showed a loss of cochlear compression. Overall, these findings suggest that some HI listeners may benefit from an enhancement of temporal envelope coding in pitch discrimination of unresolved complex tones, and that this enhancement may be also ascribed to a reduction of cochlear compression following SNHL.

© 2015 Acoustical Society of America
Reproduction of Realistic Background Noise for Testing Telecommunications Devices

A method for reproduction of sound, based on crosstalk cancellation using inverse filters, was implemented in the context of testing telecommunications devices. The effect of the regularization parameter, number of loudspeakers, type of background noise, and a technique to attenuate audible artifacts, were investigated. The quality of the reproduced sound was evaluated both objectively and subjectively with respect to the reference sounds, at points where telecommunications devices would be potentially placed around the head. The highest regularization value gave the best results, the performance was equally good when using eight or four loudspeakers, and the reproduction method was shown to be robust for different program materials. The proposed technique to reduce audible artifacts increased the perceived similarity.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Brüel & Kjær A/S
Authors: Gil Corrales, J. D. (Intern), Song, W. (Ekstern), MacDonald, E. (Intern)
Number of pages: 10
Publication date: 2015

Host publication information
Title of host publication: Proceeding of 138th Audio Engineering Society Convention
Publisher: Audio Engineering Society
Article number: 9303
Main Research Area: Technical/natural sciences
Conference: 138th International Audio Engineering Society (AES) Convention, Warsaw, Poland, 07/05/2015 - 07/05/2015
Source: PublicationPreSubmission
Source-ID: 110809009
Publication: Research - peer-review › Article in proceedings – Annual report year: 2015
This paper addresses the problem of localising multiple competing speakers in the presence of room reverberation, where sound sources can be positioned at any azimuth on the horizontal plane. To reduce the amount of front-back confusions which can occur due to the similarity of interaural time differences (ITDs) and interaural level differences (ILDs) in the front and rear hemifield, a machine hearing system is presented which combines supervised learning of binaural cues using multi-conditional training (MCT) with a head movement strategy. A systematic evaluation showed that this approach substantially reduced the amount of front-back confusions in challenging acoustic scenarios. Moreover, the system was able to generalise to a variety of different acoustic conditions not seen during training.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Sheffield
Authors: May, T. (Intern), Ma, N. (Ekstern), Brown, G. (Ekstern)
Number of pages: 5
Publication date: 2015

Signs of noise-induced neural degeneration in humans
Animal studies demonstrated that noise exposure causes a primary and selective loss of auditory-nerve fibres with low spontaneous firing rate. This neuronal impairment, if also present in humans, can be assumed to affect the processing of supra-threshold stimuli, especially in the presence of background noise, while leaving the processing of low-level stimuli unaffected. The purpose of this study was to investigate if signs of such primary neural damage from noise-exposure could also be found in noise-exposed human individuals. It was investigated: (1) If noise-exposed listeners with hearing thresholds within the “normal” range perform poorer, in terms of their speech recognition threshold in noise (SRTN), and (2) if auditory brainstem responses (ABR) reveal lower amplitude of wave I in the noise-exposed listeners. A test group of noise/music-exposed individuals and a control group were recruited. All subjects were between 18-32 years of age and had pure-tone thresholds ≤ 15 dB HL from 250-8000 Hz. Despite normal pure-tone thresholds, the noise-exposed listeners required a significantly better signal-to-noise ratio to obtain SRTN, compared to the control group. The ABR results showed significantly lower amplitude of wave I, in the left-ear, of the test group listeners. Significantly higher wave III and normal wave V were also found in the left ear of the test group listeners suggesting a compensated neural gain in the brainstem. Overall, the results from this study seem to suggest that noise exposure affects supra-threshold processing in humans before pure-tone sensitivity, raising suspicion to the hypothesis of primary neural involvement.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Rigshospitalet
Authors: Holtegaard, P. (Intern), Olsen, S. Ø. (Ekstern)
Number of pages: 8
Publication date: 2015

Simulation of realistic background noise using multiple loudspeakers

Three methods for reproduction of sound using a maximum of eight loudspeakers were investigated in the context of testing telecommunication devices. They are the four-loudspeaker-based method as described in ETSI EG 202 396-1, Higher-Order ambisonics (HOA), and a matrix inversion method. HOA optimizes the reproduced sound at a sweet spot in the center of the array with radius determined by a spherical microphone array, which is used to derive the spherical harmonics decomposition of the reference sound. The four-loudspeaker-based method equalizes the magnitude response at the ears of a head and torso simulator (HATS) for sound reproduction, while the matrix inversion method optimizes the local sound field around a few target positions. The matrix inversion method had two conditions, i.e. with or without the extra processing steps described in ETSI TS 103 224; and three sets of optimization positions were defined, i.e. the ears of the HATS, positions close to a device under test, and standardized positions as described in ETSI TS 103 224. A listening experiment was performed to evaluate the perceived quality of the reproduced sounds at the microphones close to a device under test and at the ears of the HATS. The matrix inversion method performed best when listening to the reproduced sounds at target positions used for sound-field optimization and when listening to the microphones close to the device. HOA resulted in similar perceived quality as the matrix inversion method while a large degree of perceptual degradation was observed using the four-loudspeaker-based method.

Single channel speech enhancement in the modulation domain: New insights in the modulation channel selection framework

Recently, the ideal binary mask has been introduced in the modulation domain by extending the ideal channel selection method to modulation channel selection [1]. This new method shows substantial improvement in speech intelligibility but less than its predecessor despite the higher complexity. Here, we extend the previous finding from [1] and provide a more direct comparison of binary masking in the modulation domain with binary masking in the time-frequency domain. Subjective and objective evaluations are performed and provide additional insight into modulation domain processing.
**Sound objects – Auditory objects – Musical objects**

The auditory system transforms patterns of sound energy into perceptual objects but the precise definition of an ‘auditory object’ is much debated. In the context of music listening, Pierre Schaeffer argued that ‘sound objects’ are the fundamental perceptual units in ‘musical objects’. In this paper, I review recent neurocognitive research suggesting that the auditory system is sensitive to structural information about real-world objects. Instead of focusing solely on perceptual sound features as determinants of auditory objects, I propose that real-world object properties are inherent in the organization of the auditory system and as such in music perception as well.

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Hjortkjær, J. (Intern)
Pages: 47-56
Publication date: 2015
Main Research Area: Technical/natural sciences

**Publication information**

Journal: Dansk Musikforskning Online
Issue number: SPECIAL EDITION
ISSN (Print): 1904-237X

**Sources of variability in consonant perception and their auditory correlates**

Responses obtained in consonant perception experiments typically show large variability across BmuliofthesamephoneBcident Bty(Phatakatal., 2008; Sing&Allen, 2012; Toscano&Allen, 2014). The present study investigated the influence of different potential sources of this response variability. It was discussed between source-induced variability, referring to perceptual differences caused by acoustic differences in the speech tokens and/or the masking noise tokens, and receiver-related variability, referring to perceptual differences caused by within- and across-listener uncertainty. It can be demonstrated that any physical change in the Bmulihad a measurable effect. This holds, even for slight changes in the steady-state masking-noise waveform. Furthermore, responses obtained with the BcalsBmulidiffered substantially across different normal-hearing listeners, while individual listeners were able to reproduce their responses fairly reliably. To determine how well the source-induced variability is reflected, the corresponding perceptual distances were compared to the distances between the Rof the Bmuli. Several variants of energy-based IR and amplitude-based IR were considered. The results suggest that normalized IR of the Bmuli-based representational Bon provides the best match to the perceptual data.

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Zaar, J. (Intern), Dau, T. (Intern)
Number of pages: 1
Publication date: 2015
Sources of variability in consonant perception and their auditory correlates

Responses obtained in consonant perception experiments typically show a large variability across stimuli of the same phonetic identity. The present study investigated the influence of different potential sources of this response variability. It was distinguished between source-induced variability, referring to perceptual differences caused by acoustical differences in the speech tokens and/or the masking noise tokens, and receiver-related variability, referring to perceptual differences caused by within- and across-listener uncertainty. Two experiments were conducted with normal-hearing listeners using consonant-vowel combinations (CVs) in white noise. The responses were analyzed with respect to the different sources of variability based on a measure of perceptual distance. The speech-induced variability across and within talkers and the across-listener variability were substantial and of similar magnitude. The noise-induced variability was smaller than the above-mentioned contributions but significantly larger than the amount of within-listener variability, which represented the smallest effect. To determine how well the source-induced variability is reflected in different auditory-inspired internal representations (IRs), the corresponding perceptual distances were compared to the distances between the IRs of the stimuli. Several variants of an auditory-spectrogram based IR and a modulation-spectrogram based IR were considered and the importance of the different domains for consonant perception was evaluated.

© 2015 Acoustical Society of America
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 0.848 SNIP 1.633
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 0.865 SNIP 1.647
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 0.752 SNIP 1.559
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 0.954 SNIP 1.749
Web of Science (2005): Indexed yes
Scopus rating (2004): SJR 0.77 SNIP 1.787
Web of Science (2004): Indexed yes
Scopus rating (2003): SJR 0.882 SNIP 1.712
Web of Science (2003): Indexed yes
Scopus rating (2002): SJR 0.87 SNIP 1.501
Web of Science (2002): Indexed yes
Scopus rating (2001): SJR 0.719 SNIP 1.467
Web of Science (2001): Indexed yes
Scopus rating (2000): SJR 0.621 SNIP 1.411
Web of Science (2000): Indexed yes
Scopus rating (1999): SJR 0.591 SNIP 1.319
Original language: English
DOIs:
10.1121/1.4920423
Publication: Research - peer-review › Journal article – Annual report year: 2015

Sources of variability in consonant perception of normal-hearing listeners
Responses obtained in consonant perception experiments typically show a large variability across stimuli of the same phonetic identity. The present study investigated the influence of different potential sources of this response variability. It was distinguished between source-induced variability, referring to perceptual differences caused by acoustical differences in the speech tokens and/or the masking noise tokens, and receiver-related variability, referring to perceptual differences caused by within- and across-listener uncertainty. Consonant-vowel combinations consisting of 15 consonants followed by the vowel /i/ were spoken by two talkers and presented to eight normal-hearing listeners both in quiet and in white noise at six different signal-to-noise ratios. The obtained responses were analyzed with respect to the different sources of variability using a measure of the perceptual distance between responses. The speech-induced variability across and within talkers and the across-listener variability were substantial and of similar magnitude. The noise-induced variability, obtained with time-shifted realizations of the same random process, was smaller but significantly larger than the amount of within-listener variability, which represented the smallest effect. The results have implications for the design of consonant perception experiments and provide constraints for future models of consonant perception.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Zaar, J. (Intern), Dau, T. (Intern)
Pages: 1253-1267
Publication date: 2015
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of the Acoustical Society of America
Spectral Weighting of Binaural Cues: Effect of Bandwidth and Stream Segregation

Anecdotally, normal hearing listeners can attend to a single sound source in the presence of other sound sources by forming auditory objects. This is commonly referred to as the cocktail party effect. It is known that listeners use, among others, interaural disparities in time and intensity (referred to as ITD and ILD, respectively) to localize a sound source. An open question is, however, how ITD and ILD information is integrated over frequency, and how streaming affects auditory object formation using interaural disparities. ITD weighting functions were previously derived using inverted sensitivity thresholds of narrowband signals (Stern et al., 1988). This method does not take binaural interference (McFadden and Pasanen, 1976) into account and might not be applicable to more realistic broadband signals.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Ahrens, A. (Intern), Joshi, S. N. (Intern), Epp, B. (Intern)
Number of pages: 1
Publication date: 2015
Event: Poster session presented at 38th Annual MidWinter Meeting of the Association for Research in Otolaryngology, Baltimore, MD, United States.
Main Research Area: Technical/natural sciences
Electronic versions:
Spectral_Weighting_of_Binaural_Cues.pdf
Publication: Research › Poster – Annual report year: 2015

Speech Intelligibility Evaluation for Mobile Phones.
In the development process of modern telecommunication systems, such as mobile phones, it is common practice to use computer models to objectively evaluate the transmission quality of the system, instead of time-consuming perceptual listening tests. Such models have typically focused on the quality of the transmitted speech, while little or no attention has been provided to speech intelligibility. The present study investigated to what extent three state-of-the-art speech intelligibility models could predict the intelligibility of noisy speech transmitted through mobile phones. Sentences from the Danish Dantale II speech material were mixed with three different kinds of background noise, transmitted through three different mobile phones, and recorded at the receiver via a local network simulator. The speech intelligibility of the transmitted sentences was assessed by six normal-hearing listeners and model predictions were compared to the perceptual data. Statistically significant differences between the intelligibility of the three phones were found in stationary speech-shaped noise. A good correspondence between the measured data and the predictions from one of the three models was found in all the considered conditions. Overall, the results suggest that speech intelligibility models inspired by auditory signal processing can be useful for the objective evaluation of speech transmission through mobile phones.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Jørgensen, S. (Intern), Cubick, J. (Intern), Dau, T. (Intern)
Pages: 1016 – 1025
Publication date: 2015
Main Research Area: Technical/natural sciences

Publication information
Journal: Acta Acustica United With Acustica
Volume: 101
ISSN (Print): 1610-1928
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Spektrale Gewichtung von interauralen Zeit- und Pegelunterschieden zur Lateralisierung von Breitbandsignalen

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Statistical representation of sound textures in the impaired auditory system

Many challenges exist when it comes to understanding and compensating for hearing impairment. Traditional methods, such as pure tone audiometry and speech intelligibility tests, offer insight into the deficiencies of a hearing-impaired listener, but can only partially reveal the mechanisms that underlie the hearing loss. An alternative approach is to investigate the statistical representation of sounds for hearing-impaired listeners along the auditory pathway. Using models of the auditory periphery and sound synthesis, we aimed to probe hearing impaired perception for sound textures – temporally homogenous sounds such as rain, birds, or fire. It has been suggested that sound texture perception is mediated by time-averaged statistics measured from early auditory representations (McDermott et al., 2013). Changes to early auditory processing, such as broader “peripheral” filters or reduced compression, alter the statistical representation of sound textures. We show that these changes in the statistical representation are reflected in perception, where listeners can discriminate between synthetic textures generated from normal and impaired models of the auditory periphery. Further, a simple compensation strategy was investigated to recover the perceptual qualities of a synthetic sound texture generated from an impaired model.

The influence of visual cues on auditory distance perception.

The General information section for the last document is not provided.
The role of harmonic resolvability for pitch discrimination

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Lamping, W. (Intern), Santurette, S. (Intern), MacDonald, E. (Intern)
Number of pages: 4
Publication date: 2015

The role of reverberation-related binaural cues in the externalization of speech

The perception of externalization of speech sounds was investigated with respect to the monaural and binaural cues available at the listeners' ears in a reverberant environment. Individualized binaural room impulse responses (BRIRs) were used to simulate externalized sound sources via headphones. The measured BRIRs were subsequently modified such that the proportion of the response containing binaural vs monaural information was varied. Normal-hearing listeners were presented with speech sounds convolved with such modified BRIRs. Monaural reverberation cues were found to be sufficient for the externalization of a lateral sound source. In contrast, for a frontal source, an increased amount of binaural cues from reflections was required in order to obtain well externalized sound images. It was demonstrated that the interaction between the interaural cues of the direct sound and the reverberation strongly affects the perception of externalization. An analysis of the short-term binaural cues showed that the amount of fluctuations of the binaural cues corresponded well to the externalization ratings obtained in the listening tests. The results further suggested that the precedence effect is involved in the auditory processing of the dynamic binaural cues that are utilized for externalization perception.
The role of temporal resolution in modulation-based speech segregation

This study is concerned with the challenge of automatically segregating a target speech signal from interfering background noise. A computational speech segregation system is presented which exploits logarithmically-scaled amplitude modulation spectrogram (AMS) features to distinguish between speech and noise activity on the basis of individual time-frequency (T-F) units. One important parameter of the segregation system is the window duration of the analysis-synthesis stage, which determines the lower limit of modulation frequencies that can be represented but also the temporal acuity with which the segregation system can manipulate individual T-F units. To clarify the consequences of this trade-off on modulation-based speech segregation performance, the influence of the window duration was systematically investigated.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: May, T. (Intern), Bentsen, T. (Intern), Dau, T. (Intern)
Time-efficient multidimensional threshold tracking method

Traditionally, adaptive methods have been used to reduce the time it takes to estimate psychoacoustic thresholds. However, even with adaptive methods, there are many cases where the testing time is too long to be clinically feasible, particularly when estimating thresholds as a function of another parameter, such as in temporal masking curves or when characterizing auditory filters. Here we present a new method, the “grid” method, which adaptively varies multiple parameters during each experimental run. By changing the way the parameter-response space is sampled, the method increases the proportion of experimental time spent in the vicinity of the sought-after threshold curve. The resulting increase in time-efficiency is substantial and can make some measurements clinically feasible. Thresholds from temporal masking curves obtained with the grid method are compared with those from one of the most time-efficient standard methods (single-interval up-down adaptive method of Lecluyse, 2013). Overall, individuals’ results from both methods are very highly correlated, but the grid method was an order of magnitude faster in estimating thresholds. The application of the grid method to other measurements, such as characterizing auditory filters, will also be discussed.
Using eye movements for analyzing the influence of linguistic complexity, noise, and hearing loss on sentence processing time

High linguistic complexity can reduce speech intelligibility and can increase cognitive effort. A method for detecting the latter was presented by Wendt et al. (2014) using an eye-tracking (ET) paradigm measuring increased processing time for complex sentences. This study evaluates this method and compares the ET method to electrooculography (EOG). The processing time of sentences with different linguistic complexity was measured in quiet and in modulated noise using ET and EOG simultaneously. Eleven participants with hearing impairment and five participants with normal hearing participated in the study. Processing times measured using ET and using EOG showed a correlation of 94%. Furthermore, our results confirm the findings of Wendt and colleagues, that more complex sentences show increased processing time. This study evaluated that sentence processing time can be analyzed equally well using ET and EOG. The method reveals characteristic consequences of linguistic complexity and noise on sentence processing time which can be used as an indicator of the cognitive effort during sentence comprehension.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Oldenburg
Authors: Wendt, D. (Intern), Müller, J. (Ekstern), Kollmeier, B. (Ekstern), Brand, T. (Ekstern)
Number of pages: 1
A Danish open-set speech corpus for competing-speech studies

Studies investigating speech-on-speech masking effects commonly use closed-set speech materials such as the coordinate response measure [Bolia et al. (2000). J. Acoust. Soc. Am. 107, 1065-1066]. However, these studies typically result in very low (i.e., negative) speech recognition thresholds (SRTs) when the competing speech signals are spatially separated. To achieve higher SRTs that correspond more closely to natural communication situations, an open-set, low-context, multi-talker speech corpus was developed. Three sets of 268 unique Danish sentences were created, and each set was recorded with one of three professional female talkers. The intelligibility of each sentence in the presence of speech-shaped noise was measured. For each talker, 200 approximately equally intelligible sentences were then selected and systematically distributed into 10 test lists. Test list homogeneity was assessed in a setup with a frontal target sentence and two concurrent masker sentences at ±50 degrees azimuth. For a group of 16 normal-hearing listeners and a group of 15 elderly (linearly aided) hearing-impaired listeners, overall SRTs of, respectively, +1.3 dB and +6.3 dB target-to-masker ratio were obtained. The new corpus was found to be very sensitive to inter-individual differences and produced consistent results across test lists. The corpus is publicly available.

General information

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Oticon A/S
Authors: Nielsen, J. B. (Intern), Dau, T. (Intern), Neher, T. (Ekstern)
Pages: 407-420
Publication date: 2014
Main Research Area: Technical/natural sciences

Publication information

Volume: 135
Issue number: 1
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Analyzing processing effort during sentence comprehension in quiet and in noise: Evidence from eye-fixations and pupil size

Eye-fixations can be used to investigate sentence processing and the required effort during sentence comprehension. Wendt and colleagues (Wendt et al., 2014) proposed an eye-tracking paradigm to detect time-consuming aspects during sentence processing. Participants’ eye-fixations were recorded within an audio-visual paradigm to investigate the speed of processing sentences with varying syntactic complexity. Even at high speech intelligibility levels, a reduced processing speed was measured indicating increased processing effort for complex sentences. Another measure of cognitive processing effort is served by task-evoked pupillary response. For instance, Piquard et al. (2010) showed significantly larger pupil sizes during speech comprehension for syntactically more complex object-relative sentences than for the syntactically less complex subject-relative sentence structures. Here, we compare both methods, i.e., processing speed and pupil size, as indicators for the required effort when processing sentences that differ in their level of syntactic complexity. Furthermore, an interaction of background noise and syntactic complexity is examined by analyzing...
processing effort for sentences presented in quiet and in noise. Moreover, it is investigated whether both measures provide similar or complementary information about sentence processing and the required effort.

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Oldenburg
Authors: Wendt, D. (Intern), Brand, T. (Ekstern), Kollmeier, B. (Ekstern)
Number of pages: 6
Publication date: 2014

**Host publication information**

Title of host publication: 16. Jahrestagung der Deutschen Gesellschaft für Audiologie
Main Research Area: Technical/natural sciences
Processing effort, Eye-tracking paradigm, Processing speed, Pupil size, Linguistic complexity
Source: PublicationPreSubmission
Source-ID: 100217421
Publication: Research - peer-review › Article in proceedings – Annual report year: 2014

**Assessing the effects of temporal coherence on auditory stream formation through comodulation masking release**

Recent studies of auditory streaming have suggested that repeated synchronous onsets and offsets over time, referred to as "temporal coherence," provide a strong grouping cue between acoustic components, even when they are spectrally remote. This study uses a measure of auditory stream formation, based on comodulation masking release (CMR), to assess the conditions under which a loss of temporal coherence across frequency can lead to auditory stream segregation. The measure relies on the assumption that the CMR, produced by flanking bands remote from the masker and target frequency, only occurs if the masking and flanking bands form part of the same perceptual stream. The masking and flanking bands consisted of sequences of narrowband noise bursts, and the temporal coherence between the masking and flanking bursts was manipulated in two ways: (a) By introducing a fixed temporal offset between the flanking and masking bands that varied from zero to 60 ms and (b) by presenting the flanking and masking bursts at different temporal rates, so that the asynchronies varied from burst to burst. The results showed reduced CMR in all conditions where the flanking and masking bands were temporally incoherent, in line with expectations of the temporal coherence hypothesis.

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Minnesota
Authors: Christiansen, S. K. (Intern), Oxenham, A. J. (Ekstern)
Pages: 3520–3529
Publication date: 2014
Main Research Area: Technical/natural sciences

**Publication information**

Journal: Journal of the Acoustical Society of America
Volume: 135
Issue number: 6
ISSN (Print): 0001-4966
Ratings:

- BFI (2018): BFI-level 2
- Web of Science (2018): Indexed yes
- BFI (2017): BFI-level 2
- Web of Science (2017): Indexed yes
- BFI (2016): BFI-level 2
- Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
- Web of Science (2016): Indexed yes
- BFI (2015): BFI-level 2
- Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
- Web of Science (2015): Indexed yes
- BFI (2014): BFI-level 2
- Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
- Web of Science (2014): Indexed yes
- BFI (2013): BFI-level 2
- Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
Auditory Perception of Statistically Blurred Sound Textures

Sound textures have been identified as a category of sounds which are processed by the peripheral auditory system and captured with running time-averaged statistics. Although sound textures are temporally homogeneous, they offer a listener with enough information to identify and differentiate sources. This experiment investigated the ability of the auditory system to identify statistically blurred sound textures and the perceptual relationship between sound textures. Identification performance of statistically blurred sound textures presented at a fixed blur increased over those presented as a gradual blur. The results suggest that the correct identification of sound textures is influenced by the preceding blurred stimulus. These findings draw parallels to the recognition of blurred images.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Capturing and reproducing realistic acoustic scenes for hearing research

Accurate spatial audio recordings are important for a range of applications, from the creation of realistic virtual sound environments to the evaluation of communication devices, such as hearing instruments and mobile phones. Spherical microphone arrays are particularly well-suited for capturing spatial audio in three dimensions. However, practical constraints limit the number of microphones that can be used and thus the maximum spatial resolution and frequency bandwidth that can be achieved. Further, most important sound sources are near the horizontal plane, where human spatial hearing is also most accurate. This thesis therefore investigated whether the horizontal performance of spherical microphone arrays could be improved (i) through an appropriate placement of a fixed number of transducers on the sphere, and (ii) by applying mixed-order ambisonics (MOA) processing. MOA combines higher-order ambisonics (HOA) with additional, horizontally oriented spherical harmonic functions of higher orders. Simulations of a MOA array, with a higher density of microphones near the equator, and an array with a nearly uniform distribution of microphones were compared in terms of spatial resolution and robustness. A MOA array was constructed, and some of the simulation results were validated with measurements. Results showed that for MOA, the spatial resolution was improved for horizontal sources at mid to high frequencies and the robustness to noise and measurement errors was similar to that of HOA. The properties of MOA microphone layouts and processing were investigated further by considering several order combinations. It was shown that the performance for horizontal vs. elevated sources can be adjusted by varying the order combination, but that a benefit of the higher horizontal orders can only be seen at mid to high frequencies as the need for regularization limits spatial directivity at lower frequencies. Finally, the MOA array was also evaluated in terms of sound field reconstruction error in a head-sized region. Results provided a physical validation of the functioning of the MOA microphone array and further showed that the MOA approach results in a somewhat larger “sweet area” for horizontal sources than for elevated sound sources. While the focus was on the technical evaluation of the developed MOA system, potential perceptual effects concerning MOA and microphone array recordings in general are also discussed. The system developed in this work provides new possibilities for the investigation of human perception in realistic and complex acoustic environments.

Comparison of peripheral compression estimates using auditory steady-state responses (ASSR) and distortion product otoacoustic emissions (DPOAE)

The healthy auditory system shows a compressive input/output (I/O) function as a result of healthy outer-hair cell function. Hearing impairment often leads to a decrease in sensitivity and a reduction of compression, mainly caused by loss of inner and/or outer hair cells. Compression is commonly estimated based on behavioral procedures (Plack et al., 2004), which are time consuming and rely on assumptions regarding the ability to selectively investigate cochlear processing; or on objective recordings such as otoacoustic emissions (OAEs) (Neely et al., 2003), which allow to selectively study cochlear processing but the interpretation of results for individual data is challenging. Auditory steady-state responses (ASSR) are another objective method which allows fast, reliable and frequency-specific measurements of hearing function. It is investigated here whether ASSR can be used to estimate compression along the peripheral auditory pathway. It is
hypothesized that compressive behavior is observed in normal-hearing (NH) listeners while in hearing-impaired (HI) listeners, sensitivity and compression are reduced. ASSR data are later compared to data from distortion-product otoacoustic emissions (DPOAEs) recordings. Results show compressive ASSR I/O functions for NH subjects. For HI subjects, ASSR reveal the loss of sensitivity at low stimulus levels. Growth slopes are smaller (more compressive) in ASSR than in DPOAE I/O functions.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Encina Llamas, G. (Intern), Epp, B. (Intern), Dau, T. (Intern)
Number of pages: 1
Publication date: 2014
Event: Poster session presented at 37th Annual MidWinter Meeting of the Association for Research in Otolaryngology, San Diego, CA, United States.
Main Research Area: Technical/natural sciences
Electronic versions: Comparison_of_peripheral_compression.pdf
Publication: Research › Poster – Annual report year: 2015

Computational speech segregation based on an auditory-inspired modulation analysis
A monaural speech segregation system is presented that estimates the ideal binary mask from noisy speech based on the supervised learning of amplitude modulation spectrogram (AMS) features. Instead of using linearly scaled modulation filters with constant absolute bandwidth, an auditory-inspired modulation filterbank with logarithmically scaled filters is employed. To reduce the dependency of the AMS features on the overall background noise level, a feature normalization stage is applied. In addition, a spectro-temporal integration stage is incorporated in order to exploit the context information about speech activity present in neighboring time-frequency units. In order to evaluate the generalization performance of the system to unseen acoustic conditions, the speech segregation system is trained with a limited set of low signal-to-noise ratio (SNR) conditions, but tested over a wide range of SNRs up to 20dB. A systematic evaluation of the system demonstrates that auditory-inspired modulation processing can substantially improve the mask estimation accuracy in the presence of stationary and fluctuating interferers.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: May, T. (Intern), Dau, T. (Intern)
Pages: 3350–3359
Publication date: 2014
Main Research Area: Technical/natural sciences

Publication Information
Journal: Journal of the Acoustical Society of America
Volume: 136
Issue number: 6
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Consonance perception of complex-tone dyads and chords

Sensory consonance and dissonance are perceptual attributes of musical intervals conveying pleasantness, tension, and harmony in musical phrases. For complex-tone dyads, corresponding to two musical notes played simultaneously, consonance is known to vary with the ratio in fundamental frequency (F0) between the two tones in the dyad. While such a relationship is well established for dyads, the subjective consonance of chords containing three or more simultaneous notes, that form the basis of most musical pieces, remains to be explored. The present study aimed at comparing consonance judgments for dyad/dyad, dyad/chord, and chord/chord combinations as a function of the F0 ratio between their element tones. Dyads and chords were generated by adding two or three complex tones containing 6 harmonics with equal amplitude and random phase. The base F0 of the first tone was randomly selected from an interval spanning 3 = 1/4 of a semitone centered at 440 Hz. The second tone F0 varied between 0–12 semitones above the base F0. For chords, the third tone F0 was fixed either at 5 (Perfect 4th, P4) or at 7 (Perfect 5th, P5) semitones above the base F0. Ten normal-hearing listeners were presented with all possible dyad/dyad, dyad/chord, and chord/chord combinations in random order and were asked to judge which interval was most consonant in each paired comparison. The results for dyad/dyad comparisons were consistent with earlier findings, with the unison, octave, P5, and P4 intervals being perceived as the most consonant. For dyad/chord comparisons, dyads were more consonant in the intervals around the fixed third tone. Overall, chords were not found to be more dissonant than dyads. This suggests that the hypothesis according to which consonance decreases with the amount...
of interaction between present harmonics, arguing for a potential role of frequency selectivity for consonance perception of dyads, might not hold for chords.

**General information**
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Technical University of Denmark
Authors: Rasmussen, M. (Ekstern), Santurette, S. (Intern), MacDonald, E. (Intern)
Number of pages: 4
Publication date: 2014

**Host publication information**
Title of host publication: Proceedings of Forum Acusticum
BFI conference series: Forum Acusticum (5010988)
Main Research Area: Technical/natural sciences
Electronic versions:
Rasmussen2014.pdf
Source: PublicationPreSubmission
Source-ID: 99975576
Publication: Research - peer-review › Article in proceedings – Annual report year: 2014

**Consonant confusions in frozen and random white noise**

**General information**
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Zaa, J. (Intern), Jørgensen, S. (Intern), Dau, T. (Intern)
Number of pages: 1
Publication date: 2014
Event: Poster session presented at 6th Workshop on Speech in Noise, Marseille, France.
Main Research Area: Technical/natural sciences
Electronic versions:
ConsonantConfusions_poster_JZaar_SpIN2014.pdf
Source: PublicationPreSubmission
Source-ID: 102826367
Publication: Research - peer-review › Poster – Annual report year: 2014

**Consonant confusions in frozen and random white noise**

**General information**
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Zaa, J. (Intern), Jørgensen, S. (Intern), Dau, T. (Intern)
Number of pages: 1
Publication date: 2014
Event: Poster session presented at Concepts and computational models of robust bottom-up signal encoding, Copenhagen, Denmark.
Main Research Area: Technical/natural sciences
Electronic versions:
ConsonantConfusions_poster_JZaar_SpIN2014.pdf

**Bibliographical note**
Poster presented at the INSPIRE Copenhagen Winter School, the topic of which was: "Concepts and computational models of robust bottom-up encoding”
Publication: Research - peer-review › Poster – Annual report year: 2014

**Distortion-product otoacoustic emission reflection-component delays and cochlear tuning: Estimates from across the human lifespan**
A consistent relationship between reflection-emission delay and cochlear tuning has been demonstrated in a variety of mammalian species, as predicted by filter theory and models of otoacoustic emission (OAE) generation. As a step toward the goal of studying cochlear tuning throughout the human lifespan, this paper exploits the relationship and explores two strategies for estimating delay trends—energy weighting and peak picking—both of which emphasize data at the peaks of the magnitude fine structure. Distortion product otoacoustic emissions (DPOAEs) at 2f1f2 were recorded, and their reflection components were extracted in 184 subjects ranging in age from prematurely born neonates to elderly adults.
DPOAEs were measured from 0.5–4 kHz in all age groups and extended to 8 kHz in young adults. Delay trends were effectively estimated using either energy weighting or peak picking, with the former method yielding slightly shorter delays and the latter somewhat smaller confidence intervals. Delay and tuning estimates from young adults roughly match those obtained from SFOAEs. Although the match is imperfect, reflection-component delays showed the expected bend (apical-basal transition) near 1 kHz, consistent with a break in cochlear scaling. Consistent with other measures of tuning, the term newborn group showed the longest delays and sharpest tuning over much of the frequency range.

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, House Research Institute, Harvard Medical School
Authors: Abdala, C. (Ekstern), Guérit, F. (Intern), Luo, P. (Ekstern), Shera, C. A. (Ekstern)
Pages: 1950–1958
Publication date: 2014
Main Research Area: Technical/natural sciences

**Publication information**

Journal: Journal of the Acoustical Society of America
Volume: 135
Issue number: 4
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 0.848 SNIP 1.633
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 0.865 SNIP 1.647
Web of Science (2007): Indexed yes
Effects of spontaneous otoacoustic emissions on pure-tone frequency difference limens

Pure-tone frequency difference limens (FDLs) have been shown to vary in the vicinity of spontaneous otoacoustic emissions (SOAEs). As lower FDLs have been observed near SOAEs when measured ipsi- and contralaterally to the emission ear, it has been proposed that prolonged ongoing stimulation of nerve cells tuned to the SOAE frequency could lead to a central oversensitivity to that frequency, hence a better frequency-discrimination ability. However, it is also known that tones close in frequency to an SOAE can "entrain" the emission to oscillate at their own frequency. This may instead explain the variations in FDL near SOAE frequencies as arising from peripheral interactions between SOAEs and external tones in the cochlea. To test these two hypotheses, SOAE entrainment patterns and FDLs were recorded in seven subjects with an ipsilateral SOAE and no neighboring contralateral SOAE. Ipsilateral FDLs were lowest in the SOAE entrainment region and worsened significantly when beating between the external tone and SOAE occurred. FDLs remained unaffected in the non-emission ear and did not alter with continuous ipsilateral or contralateral presentation of a pure tone aimed at emulating an SOAE. These findings suggest a mechanical rather than neural origin for the variations in FDL near SOAE frequencies.
Effects of tonotopicity, adaptation, modulation tuning, and temporal coherence in "primitive" auditory stream segregation

The perceptual organization of two-tone sequences into auditory streams was investigated using a modeling framework consisting of an auditory pre-processing front end [Dau et al., J. Acoust. Soc. Am. 102, 2892–2905 (1997)] combined with a temporal coherence-analysis back end [Elhilali et al., Neuron 61, 317–329 (2009)]. Two experimental paradigms were considered: (i) Stream segregation as a function of tone repetition time (TRT) and frequency separation (DF) and (ii) grouping of distant spectral components based on onset/offset synchrony. The simulated and experimental results of the
The present study supported the hypothesis that forward masking enhances the ability to perceptually segregate spectrally close tone sequences. Furthermore, the modeling suggested that effects of neural adaptation and processing through modulation-frequency selective filters may enhance the sensitivity to onset asynchrony of spectral components, facilitating the listeners' ability to segregate temporally overlapping sounds into separate auditory objects. Overall, the modeling framework may be useful to study the contributions of bottom-up auditory features on “primitive” grouping, also in more complex acoustic scenarios than those considered here.
Electrophysiological assessment of audiovisual integration in speech perception

Speech perception integrates signal from ear and eye. This is witnessed by a wide range of audiovisual integration effects, such as ventriloquism and the McGurk illusion. Some behavioral evidence suggest that audiovisual integration of specific aspects is special for speech perception. However, our knowledge of such bimodal integration would be strengthened if the phenomena could be investigated by objective, neutrally based methods. One key question of the present work is if perceptual processing of audiovisual speech can be gauged with a specific signature of neurophysiological activity, the mismatch negativity response (MMN).

MMN has the property of being evoked when an acoustic stimulus deviates from a learned pattern of stimuli. In three experimental studies, this effect is utilized to track when a coinciding visual signal alters auditory speech perception. Visual speech emanates from the face of the talker. Perception of faces and of speech shares the trait, that they are learned from infancy and seemingly specialized behaviorally and neurally. Due to this, speech and face encoding functions quasi-automatically and with high efficiency. However, perhaps owing to our long experience with human faces, which all are variations on a relatively constrained space of features, face perception is sensitive to manipulations of the structure of the face, the relation between its segments, and the properties of the segments. Does this sensitivity alter the influence of visual speech on the auditory speech percept? In two experiments, which both combine behavioral and neurophysiological measures, an uncovering of the relation between perception of faces and of audiovisual integration is attempted. Behavioral findings suggest a strong effect of face perception, whereas the MMN results are less clear.

Another interesting property of speech perception is that it is relatively tolerant towards temporal shifts between acoustic and visual speech signals. Here, behavioral studies report that perception of speech exhibits far greater temporal tolerance than towards non-speech stimuli. Current findings on neural correlates of this tolerance, however, are few and limited. Here, a novel experimental MMN paradigm is used in effort to shed light on integration asynchronous audiovisual speech. Based on individual behavioral estimates of temporal windows of tolerance, we ask if the MMN signal can be evoked at different points within and outside this window. Behavioral findings match earlier behavioral studies, whereas the MMN findings are ambiguous.

In conclusion, the work presented here sheds light onto two important aspects of speech perception. It also presents important methodological conclusions on the use of MMN as a neural marker of audiovisual integration.
Eliminating transducer distortion in acoustic measurements
This paper investigates the influence of nonlinear components that contaminate the linear response of acoustic transducer, and presents a method for eliminating the influence of nonlinearities in acoustic measurements. The method is evaluated on simulated as well as experimental data, and is shown to perform well even in noisy conditions. The limitations of the Total Harmonic Distortion, THD, measure is discussed and a new distortion measure, Total Distortion Ratio, TDR, which more accurately describes the amount of nonlinear power in the measured signal, is proposed.

Ellipsoidal reflector for measuring oto-acoustic emissions
A truncated prolate ellipsoidal reflector having the ear canal of a listener at one focal point and large-diaphragm low-noise microphone at the other focal point is proposed for free-field recordings of oto-acoustic emissions. A prototype reflector consisting of three pieces is presented, which enables measuring the response of the system with different truncations. The response of the system is measured with a miniature loud-speaker, and proof-of-concept measurements of oto-acoustic emissions are presented. The effect of truncation and other physical parameters to the performance of the system are discussed.

Exploring the physical correlates of consonant recognition and confusions
General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Zaar, J. (Intern), Jørgensen, S. (Intern), Dau, T. (Intern)
Generalization of Supervised Learning for Binary Mask Estimation

This paper addresses the problem of speech segregation by estimating the ideal binary mask (IBM) from noisy speech. Two methods will be compared, one supervised learning approach that incorporates a priori knowledge about the feature distribution observed during training. The second method solely relies on a frame-based speech presence probability (SPP) estimation, and therefore, does not depend on the acoustic condition seen during training. We investigate the influence of mismatches between the acoustic conditions used for training and testing on the IBM estimation performance and discuss the advantages of both approaches.

Human Sound Externalization in Reverberant Environments

In everyday environments, listeners perceive sound sources as externalized. In listening conditions where the spatial cues that are relevant for externalization are not represented correctly, such as when listening through headphones or hearing aids, a degraded perception of externalization may occur. In this thesis, the spatial cues that arise from a combined effect of filtering due to the head, torso, and pinna and the acoustic environment were analysed and the impact of such cues for the perception of externalization in different frequency regions was investigated. Distant sound sources were simulated via headphones using individualized binaural room impulse responses (BRIRs).

An investigation of the influence of spectral content of a sound source on externalization showed that effective externalization cues are present across the entire frequency range. The fluctuation of interaural level differences (ILDs) that occurs in reverberant environments was altered via modifications of the signal envelope in the left and right ear. It was found that the dynamic ILDs had an effect on externalization for broadband and highpass filtered speech, while no effect was found for lowpass filtered speech. Moreover, the compression of low frequency ITD fluctuations did not influence externalization.

Further, the influence of binaural and monaural cues from reverberation was investigated. It was found that monaural reverberation cues were sufficient for the externalization of a lateral source, whereas a frontal source required an increased amount of binaural cues from reflections in order to attain convincingly externalized sound images. It was concluded that the disparity in the interaural cues of the direct sound and the reverberation was important as the interaction of the two played a role for the perception of externalization. Moreover, similar binaural effects of reverberation were found at low- and high frequencies.

The performance of a multi-microphone noise reduction algorithm designed to preserve binaural cues in hearing aids was investigated in conjunction with a voice activity detector (VAD) for noise estimation. Intelligibility weighted improvements in signal-to-noise ratio (SNR) of 6 dB and 18 dB were found for diffuse multi-talker babble noise and speech shaped directional noise, respectively, at input SNRs close to the speech reception threshold (SRT) of hearing impaired listeners. Overall, this work contributes to the understanding of the auditory processing of spatial cues that are important for externalization in reverberant environments and may have implications for hearing instrument signal processing.
Influence of acoustic complexity on spatial release from masking and lateralization.

In realistic listening scenarios, humans are extremely skillful in following one particular talker even in the presence of many others (i.e., the cocktail party effect). One aspect of this is the ability of a listener to make use of the spatial separation between different sound sources. In a complex acoustic scene, as interferers are moved away from the spatial position of the target, speech intelligibility (SI) increases, often referred to as spatial release from masking (SRM). This benefit is largely based on the listeners’ ability to make use of interaural level differences (ILD) and interaural time differences (ITD), which vary with the source location. While many studies have explored SRM, few have investigated the effects of overall number and spatial distribution of interferers while controlling for monaural masking effects. In the present study, speech reception thresholds (SRTs) and lateralization thresholds were measured over headphones in babble noise conditions consisting of 2, 4, 8 and 12 talkers. The perceived locations of the signal (female voice) and individual maskers (male, time-reversed voices) were steered separately either to the left or to the right using 680 sec ITDs. For a fixed number of maskers, the distribution of interfering talkers (i.e., co-located or separated from the target) was varied. Thus, for all conditions with the same number of maskers, the monaural SNR was held constant regardless of the perceived spatial distribution. The performance between the speech and lateralization tasks was highly correlated. No substantial SRM occurred while one or more maskers were co-located with the target. Interestingly, spatially shifting the last 2 co-located interferers resulted in the same SRM, independent of the overall number of maskers. The same finding was found for the last 4 co-located maskers. The results suggest that SRM is independent of the overall number of interfering talkers.

Influence of High-Frequency Audibility on Distance Perception

In realistic listening scenarios, humans are extremely skillful in following one particular talker even in the presence of many others (i.e., the cocktail party effect). One aspect of this is the ability of a listener to make use of the spatial separation between different sound sources. In a complex acoustic scene, as interferers are moved away from the spatial position of the target, speech intelligibility (SI) increases, often referred to as spatial release from masking (SRM). This benefit is largely based on the listeners’ ability to make use of interaural level differences (ILD) and interaural time differences (ITD), which vary with the source location. While many studies have explored SRM, few have investigated the effects of overall number and spatial distribution of interferers while controlling for monaural masking effects. In the present study, speech reception thresholds (SRTs) and lateralization thresholds were measured over headphones in babble noise conditions consisting of 2, 4, 8 and 12 talkers. The perceived locations of the signal (female voice) and individual maskers (male, time-reversed voices) were steered separately either to the left or to the right using 680 sec ITDs. For a fixed number of maskers, the distribution of interfering talkers (i.e., co-located or separated from the target) was varied. Thus, for all conditions with the same number of maskers, the monaural SNR was held constant regardless of the perceived spatial distribution. The performance between the speech and lateralization tasks was highly correlated. No substantial SRM occurred while one or more maskers were co-located with the target. Interestingly, spatially shifting the last 2 co-located interferers resulted in the same SRM, independent of the overall number of maskers. The same finding was found for the last 4 co-located maskers. The results suggest that SRM is independent of the overall number of interfering talkers.

Influence of High-Frequency Audibility on Distance Perception

In realistic listening scenarios, humans are extremely skillful in following one particular talker even in the presence of many others (i.e., the cocktail party effect). One aspect of this is the ability of a listener to make use of the spatial separation between different sound sources. In a complex acoustic scene, as interferers are moved away from the spatial position of the target, speech intelligibility (SI) increases, often referred to as spatial release from masking (SRM). This benefit is largely based on the listeners’ ability to make use of interaural level differences (ILD) and interaural time differences (ITD), which vary with the source location. While many studies have explored SRM, few have investigated the effects of overall number and spatial distribution of interferers while controlling for monaural masking effects. In the present study, speech reception thresholds (SRTs) and lateralization thresholds were measured over headphones in babble noise conditions consisting of 2, 4, 8 and 12 talkers. The perceived locations of the signal (female voice) and individual maskers (male, time-reversed voices) were steered separately either to the left or to the right using 680 sec ITDs. For a fixed number of maskers, the distribution of interfering talkers (i.e., co-located or separated from the target) was varied. Thus, for all conditions with the same number of maskers, the monaural SNR was held constant regardless of the perceived spatial distribution. The performance between the speech and lateralization tasks was highly correlated. No substantial SRM occurred while one or more maskers were co-located with the target. Interestingly, spatially shifting the last 2 co-located interferers resulted in the same SRM, independent of the overall number of maskers. The same finding was found for the last 4 co-located maskers. The results suggest that SRM is independent of the overall number of interfering talkers.

Influence of High-Frequency Audibility on Distance Perception

In realistic listening scenarios, humans are extremely skillful in following one particular talker even in the presence of many others (i.e., the cocktail party effect). One aspect of this is the ability of a listener to make use of the spatial separation between different sound sources. In a complex acoustic scene, as interferers are moved away from the spatial position of the target, speech intelligibility (SI) increases, often referred to as spatial release from masking (SRM). This benefit is largely based on the listeners’ ability to make use of interaural level differences (ILD) and interaural time differences (ITD), which vary with the source location. While many studies have explored SRM, few have investigated the effects of overall number and spatial distribution of interferers while controlling for monaural masking effects. In the present study, speech reception thresholds (SRTs) and lateralization thresholds were measured over headphones in babble noise conditions consisting of 2, 4, 8 and 12 talkers. The perceived locations of the signal (female voice) and individual maskers (male, time-reversed voices) were steered separately either to the left or to the right using 680 sec ITDs. For a fixed number of maskers, the distribution of interfering talkers (i.e., co-located or separated from the target) was varied. Thus, for all conditions with the same number of maskers, the monaural SNR was held constant regardless of the perceived spatial distribution. The performance between the speech and lateralization tasks was highly correlated. No substantial SRM occurred while one or more maskers were co-located with the target. Interestingly, spatially shifting the last 2 co-located interferers resulted in the same SRM, independent of the overall number of maskers. The same finding was found for the last 4 co-located maskers. The results suggest that SRM is independent of the overall number of interfering talkers.
Influence of high-frequency audibility on the perceived distance of sounds
When listening in natural environments, normal-hearing (NH) listeners usually perceive sounds out-side their head, i.e.,
externalized. Sounds perceived inside the head are called internalized. Hearing-impaired (HI) listeners have been
reported to externalize sounds less accurately than NH listeners. In a study by Boyd et al. (2012), the average
externalization ratings of NH listeners dropped and matched those of HI listeners when the signals were lowpass-filtered
at 6.5 kHz. This suggested that reduced high-frequency audibility might cause a reduced externalization in HI listeners.
The present study aimed at clarifying whether the perceived distance of sounds in HI listeners differs from NH data as well
and, if so, whether distance-rating performance improves when reduced audibility is compensated for by amplification.
Individual binaural room impulse responses (BRIRs) were measured for nine different loudspeaker distances. NH and HI
listeners were asked to rate the perceived distance of processed speech samples in a MUSHRA-like test paradigm
according to optical markers placed in the same workshop room where the BRIR measurements were performed. NH
listeners rated the distance of unfiltered and lowpass-filtered speech, and HI listeners that of unfiltered speech and of
speech amplified to compensate for their audibility loss. The results for NH listeners showed no systematic effect of
lowpass-filtering the stimuli at 2 kHz or 6 kHz on distance ratings and the measured distance curves were much steeper
than those found in the literature. Preliminary results for three HI listeners showed large inter-subject variability, but as a
tendency, the distance rating seemed to vary with the energy content of the signal rather than the bandwidth, indicating
that loudness might be a strong contributor to distance perception in HI listeners.

Influence of impedance phase angle on sound pressures and reverberation times in a rectangular room
In most room acoustic predictions, phase shift on reflection has been overlooked. This study aims to quantify the effects of
the surface impedance phase angle of the boundary surfaces on room acoustic conditions. As a preliminary attempt, a
medium-sized rectangular room is simulated by a phased beam tracing model, after verifying it numerically against
boundary element simulations. First, the absorption characteristic of the boundary surfaces varies uniformly from 0.2 to
0.8, but with various impedance phase angles. Second, typical non-uniform cases having hard walls and floor, but with an
absorptive ceiling are investigated. The zero phase angle, which has commonly been assumed in practice, is regarded as
reference and differences in the sound pressure level and early decay time from the reference are quantified. As
expected, larger differences in the room acoustic parameters are found for larger impedance phase angles. Additionally,
binaural impulse responses are compared in a listening test for the uniform absorption cases, revealing that non-zero
impedance phase angle cases can be perceptually different from the reference condition in terms of reverberance
perception. For the non-uniform settings, the change in the impedance phase angle of the ceiling does not affect the
acoustic conditions significantly.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Cubick, J. (Intern), Santurette, S. (Intern), Dau, T. (Intern)
Publication date: 2014

Host publication information
Title of host publication: Proceedings of Forum Acusticum
BFI conference series: Forum Acusticum (5010988)
Main Research Area: Technical/natural sciences
Electronic versions:
Cubick_ForumAcusticum2014.pdf
Source: PublicationPreSubmission
Source-ID: 99976058
Publication: Research - peer-review › Article in proceedings – Annual report year: 2014

Influence of high-frequency audibility on the perceived distance of sounds
When listening in natural environments, normal-hearing (NH) listeners usually perceive sounds out-side their head, i.e.,
externalized. Sounds perceived inside the head are called internalized. Hearing-impaired (HI) listeners have been
reported to externalize sounds less accurately than NH listeners. In a study by Boyd et al. (2012), the average
externalization ratings of NH listeners dropped and matched those of HI listeners when the signals were lowpass-filtered
at 6.5 kHz. This suggested that reduced high-frequency audibility might cause a reduced externalization in HI listeners.
The present study aimed at clarifying whether the perceived distance of sounds in HI listeners differs from NH data as well
and, if so, whether distance-rating performance improves when reduced audibility is compensated for by amplification.
Individual binaural room impulse responses (BRIRs) were measured for nine different loudspeaker distances. NH and HI
listeners were asked to rate the perceived distance of processed speech samples in a MUSHRA-like test paradigm
according to optical markers placed in the same workshop room where the BRIR measurements were performed. NH
listeners rated the distance of unfiltered and lowpass-filtered speech, and HI listeners that of unfiltered speech and of
speech amplified to compensate for their audibility loss. The results for NH listeners showed no systematic effect of
lowpass-filtering the stimuli at 2 kHz or 6 kHz on distance ratings and the measured distance curves were much steeper
than those found in the literature. Preliminary results for three HI listeners showed large inter-subject variability, but as a
tendency, the distance rating seemed to vary with the energy content of the signal rather than the bandwidth, indicating
that loudness might be a strong contributor to distance perception in HI listeners.

Influence of impedance phase angle on sound pressures and reverberation times in a rectangular room
In most room acoustic predictions, phase shift on reflection has been overlooked. This study aims to quantify the effects of
the surface impedance phase angle of the boundary surfaces on room acoustic conditions. As a preliminary attempt, a
medium-sized rectangular room is simulated by a phased beam tracing model, after verifying it numerically against
boundary element simulations. First, the absorption characteristic of the boundary surfaces varies uniformly from 0.2 to
0.8, but with various impedance phase angles. Second, typical non-uniform cases having hard walls and floor, but with an
absorptive ceiling are investigated. The zero phase angle, which has commonly been assumed in practice, is regarded as
reference and differences in the sound pressure level and early decay time from the reference are quantified. As
expected, larger differences in the room acoustic parameters are found for larger impedance phase angles. Additionally,
binaural impulse responses are compared in a listening test for the uniform absorption cases, revealing that non-zero
impedance phase angle cases can be perceptually different from the reference condition in terms of reverberance
perception. For the non-uniform settings, the change in the impedance phase angle of the ceiling does not affect the
acoustic conditions significantly.

General information
State: Published
Organisations: Department of Electrical Engineering, Acoustic Technology, Hearing Systems, University of Sydney, Korea
Advanced Institute of Science & Technology
Authors: Jeong, C. (Intern), Lee, D. (Ekstern), Santurette, S. (Intern), Ih, J. (Ekstern)
Pages: 712–723
Publication date: 2014
Main Research Area: Technical/natural sciences
Interaction of Working Memory, Compressor Speed and Background Noise Characteristics

Previous studies have shown that individuals with poor working memory perform worse in speech recognition tests when fast compressor release time is applied. However, it is not clear why this effect occurs only when modulations are present in the background noise. This relationship is important to understand because the majority of everyday listening situations involve modulated noise. The investigation was carried out by testing two groups of older adults with similar degrees of mild-to-moderate sensorineural loss but different working memory abilities. The two groups were tested in their ability to understand speech signals presented with no modulation in the background noise, processed with slow and fast compression. The extent of background noise modulation was varied. Results suggest that the combined effect of short compressor release times, low working memory capacity, and glimpsing due to presence of amplitude modulation results in poorer speech recognition performance. There was no interaction between working memory and different noise backgrounds, with the poor working memory group demonstrating susceptibility to fast compressor in the background noise conditions.

Interaural Place-Mismatch Estimation With Two-Formant Vowels in Unilateral Cochlear Implant Users

Background: For patients with one cochlear implant (CI) and residual hearing in the opposite ear, a default frequency-to-electrode map is typically used despite large individual differences in electrode-array insertion depth. This non-individualized fitting rationale might partly explain the variability in long-term speech-reception benefit among CI users. Knowledge about the electrode-array location is thus crucial for adequate fitting. Although electrode location can theoretically be determined from CT scans, these are often unavailable in audiological practice. Moreover, existing behavioral procedures such as interaural pitch-matching are rather tedious and time-consuming. Here, an alternative method using two-formant vowels was developed and tested. Methods Eight normal-hearing (NH) listeners were presented synthesized two-formant vowels embedded between consonants /t/ and /k/, with first-formant frequencies (F1) at 250 and 400 Hz and second-formant frequencies (F2) between 600 and 2200 Hz. F1 was presented unaltered to the left ear, while F2 was presented to the right ear via a vocoder system simulating 3 different CI insertion depths. In each condition, the listeners indicated in a forced-choice task which of 6 vowels they perceived for different [F1, F2] combinations. Ten CI users (5 bimodal and 5 single-sided deaf) performed the same task for F1 presented acoustically to the non-CI ear and F2 presented either acoustically to the same ear or electrically to the CI ear. Results After some training, all NH listeners were able to fuse the unaltered F1 and vocoded F2 into a single vowel percept, and vowel distributions could be reliably derived in 7 listeners. Vocoder simulations of reduced CI insertion depth led to clear vowel-distribution shifts in these listeners. However, these shifts were overall smaller than their theoretical value, with high across-subject variability. Vowel distributions could be derived for all CI users in the monaural acoustic condition, indicating an ability to perform the task reliably. Despite this, large individual differences were observed for dichotic bimodal stimulation, with listeners showing either basal or apical shifts, or generally-poor vowel discrimination. Conclusion The two-formant-vowel method is a fast and clinic-friendly candidate to derive interaural place mismatches from a simple vowel-recognition task. However, it remains unclear whether the measured “vowel spaces” in CI users are directly related to insertion depth, and whether they are influenced by the ability to fuse acoustic and electric stimuli or habituation to the CI. The comparison of the present results to CT-scan and speech-intelligibility data in the same listeners will shed light on the validity of the proposed method.
Investigating Interaural Frequency-Place Mismatches via Bimodal Vowel Integration

For patients having residual hearing in one ear and a cochlear implant (CI) in the opposite ear, interaural place-pitch mismatches might be partly responsible for the large variability in individual benefit. Behavioral pitch-matching between the two ears has been suggested as a way to individualize the fitting of the frequency-to-electrode map but is rather tedious and unreliable. Here, an alternative method using two-formant vowels was developed and tested. The interaural spectral shift was inferred by comparing vowel spaces, measured by presenting the first formant (F1) to the nonimplanted ear and the second (F2) on either side. The method was first evaluated with eight normal-hearing listeners and vocoder simulations, before being tested with 11 CI users. Average vowel distributions across subjects showed a similar pattern when presenting F2 on either side, suggesting acclimatization to the frequency map. However, individual vowel spaces with F2 presented to the implant did not allow a reliable estimation of the interaural mismatch. These results suggest that interaural frequency-place mismatches can be derived from such vowel spaces. However, the method remains limited by difficulties in bimodal fusion of the two formants.
Maximum Acceptable Vibrato Excursion as a Function of Vibrato Rate in Musicians and Non-musicians

Human vibrato is mainly characterized by two parameters: vibrato extent and vibrato rate. These parameters have been found to exhibit an interaction both in physical recordings of singers’ voices and in listener’s preference ratings. This study was concerned with the way in which the maximum acceptable vibrato excursion varies as a function of vibrato rate in normal-hearing (NH) musicians and non-musicians. Eight NH musicians and six non-musicians adjusted the maximum vibrato excursion of a synthesized vowel for vibrato rates between 3 and 8 Hz. Individual thresholds varied across vibrato rate and, in most listeners, exhibited a peak at medium vibrato rates (5–7 Hz). Large across-subject variability was observed, and no significant effect of musical experience was found. Overall, most listeners were not solely sensitive to the vibrato excursion and there was a listener-dependent rate for which larger vibrato excursions were favored. The observed interaction between maximum excursion thresholds and vibrato rate may be due to the listeners’ judgments relying on cues provided by the rate of frequency changes (RFC) rather than excursion per se. Further studies are needed to evaluate the contribution of the RFC to vibrato perception and the possible effects of age and hearing impairment.

Modeling auditory-nerve responses to electrical stimulation

Cochlear implants (CI) directly stimulate the auditory nerve (AN), bypassing the mechano-electricaltransduction in the inner ear. Trains of biphasic, charge-balanced pulses (anodic and cathodic) areused as stimuli to avoid damage of the tissue. The pulses of either polarity are capable of producing action potentials (AP) whereby the sites of initiation of the AP differ for the two polarities. A cathodic pulse triggers an AP in the peripheral axon, whereas an anodic pulse triggers an AP in the centralaxon. The latency difference between the APs initiated at the different sites is about 200μs, whichis large enough to affect the temporal coding of sounds and hence, potentially, the communication abilities of the CI listener. In the present study, two recently proposed models of electric stimulationof the AN [1, 2, 3] were considered in terms of their efficacy to predict the spike timing for anodic andcathodic stimulation of the AN of cat [4]. The models’ responses to the electrical pulses of variousshapes [5] were also analyzed. It was found that, while the models can account for the ring rates irrespection to various biphasic pulse shapes, they fail to correctly describe the timing of AP in response to monophasic pulses. Strategies for improving the model performance with respect to correct AP timing are discussed.
Eventually, a model that is able to account for correct spike timing in electric hearing will be useful for objective evaluation and improvement of CI stimulation strategies.

**General information**
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Joshi, S. N. (Intern), Dau, T. (Intern), Epp, B. (Intern)
Number of pages: 5
Publication date: 2014

**Host publication information**
Title of host publication: Proceedings of Forum Acusticum
BFI conference series: Forum Acusticum (5010988)
Main Research Area: Technical/natural sciences
Source: PublicationPreSubmission
Source-ID: 102863634
Publication: Research - peer-review › Article in proceedings – Annual report year: 2014

**Modeling auditory-nerve responses to electrical stimulation**
Cochlear implants (CI) directly stimulate the auditory nerve (AN), bypassing the mechano-electrical transduction in the inner ear. Trains of biphasic, charge balanced pulses (anodic and cathodic) are used as stimuli to avoid damage of the tissue. The pulses of either polarity are capable of producing action potentials (AP) whereby the sites of initiation of the AP differ for the two polarities. A cathodic pulse triggers an AP in the peripheral axon, whereas an anodic pulse triggers an AP in the central axon. The latency difference between the APs initiated at the different sites is about 200μs, which is large enough to affect the temporal coding of sounds and hence, potentially, the communication abilities of the CI listener. In the present study, two recently proposed models of electric stimulation of the AN [1,2] were considered in terms of their efficacy to predict the spike timing for anodic and cathodic stimulation of the AN of cat [3]. The models’ responses to the electrical pulses of various shapes [4,5,6] were also analyzed. It was found that, while the models can account for the firing rates in response to various biphasic pulse shapes, they fail to correctly describe the timing of AP in response to monophasic pulses. Strategies for improving the model performance with respect to correct AP timing are discussed.

**General information**
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Joshi, S. N. (Intern), Dau, T. (Intern), Epp, B. (Intern)
Publication date: 2014
Event: Poster session presented at 7th Forum Acusticum, Krakow, Poland.
Main Research Area: Technical/natural sciences
Source: PublicationPreSubmission
Source-ID: 102863634
Publication: Research - peer-review › Poster – Annual report year: 2014

**Modeling consonant perception in normal-hearing listeners**
Speech perception is often studied in terms of natural meaningful speech, i.e., by measuring the intelligibility of a given set of single words or full sentences. However, when trying to understand how background noise, various sorts of transmission channels (e.g., mobile phones) or hearing impairment affect speech perception, it is advantageous to study the impact of these factors on the perception of the fundamental building blocks of speech. Non-sense syllables consisting of consonants and vowels have thus typically been presented to listeners in masking noise at various signal-to-noise ratios (SNRs). This "microscopic" approach allows for a detailed investigation of the mapping between the acoustical stimulus and the resulting percept. In the present study, an experiment with 8 native Danish normal-hearing listeners was conducted. The listeners were presented with consonant-vowel combinations (CVs) in quiet and at 6 different SNRs in white noise. The responses were analyzed in terms of recognition scores and consonant confusions. Inspired by models designed for long-term ("macroscopic") speech intelligibility prediction, two modeling concepts were considered to describe the consonant perception data: (i) an audibility-based approach, which corresponds to the Articulation Index (AI), and (ii) a modulation-masking based approach, as reflected in the speech-based Envelope Power Spectrum Model (sEPSM). For both models, the internal representations of the same stimuli as used in the experiment were calculated and fed into a template-matching back end. Using the experimental data as a reference, the resulting predictions of the two modeling approaches were compared and their respective suitability for the prediction of consonant perception was evaluated.

**General information**
State: Published
Modeling speech intelligibility based on the signal-to-noise envelope power ratio

The intelligibility of speech depends on factors related to the auditory processes involved in sound perception as well as on the acoustic properties of the sound entering the ear. However, a clear understanding of speech perception in complex acoustic conditions and, in particular, a quantitative description of the involved auditory processes provides a major challenge in speech and hearing research. This thesis presents a computational model that attempts to predict the speech intelligibility obtained by normal-hearing listeners in various adverse conditions. The model combines the concept of modulation frequency selectivity in the auditory processing of sound with a decision metric for intelligibility that is based on the signal-to-noise envelope power ratio (SNRenv). The proposed speech-based envelope power spectrum model (sEPSM) is demonstrated to account for the effects of stationary background noise, reverberation and noise reduction processing on speech intelligibility, indicating that the model is more general than traditional modeling approaches. Moreover, the model accounts for phase distortions when it includes a mechanism that evaluates the variation of envelope power across (audio) frequency. However, because the SNRenv is based on the long-term average envelope power, the model cannot account for the greater intelligibility typically observed in fluctuating noise compared to stationary noise. To overcome this limitation, a multi-resolution version of the sEPSM is presented where the SNRenv is estimated in temporal segments with a modulation-filter dependent duration. This multi-resolution approach effectively extends the applicability of the sEPSM to account for conditions with fluctuating interferers, while keeping its predictive power in the conditions with noisy speech distorted by reverberation or spectral subtraction. The relationship between the SNRenv based decision-metric and psychoacoustic speech intelligibility is further evaluated by generating stimuli with different SNRenv but the same overall power SNR. The results from the corresponding psychoacoustic data generally support the above relationship. However, the model is limited in conditions with manipulated clean speech since it does not account for the accompanied effects of speech distortions on intelligibility. The value of the sEPSM is further considered in conditions with noisy speech transmitted through three commercially available mobile phones. The model successfully accounts for the performance across the phones in conditions with a stationary speech-shaped background noise, whereas deviations were observed in conditions with “Traffic” and “Pub” noise. Overall, the results of this thesis support the hypothesis that the SNRenv is a powerful objective metric for speech intelligibility prediction. Moreover, the findings suggest that the concept of modulation-frequency selective processing in the auditory system is crucial for human speech perception.
Modeling Speech Intelligibility in Hearing Impaired Listeners

Models of speech intelligibility (SI) have a long history, starting with the articulation index (AI, [17]), followed by the SI index (SI I, [18]) and the speech transmission index (STI, [7]), to only name a few. However, these models fail to accurately predict SI with nonlinearly processed noisy speech, e.g. phase jitter or spectral subtraction. Recent studies predict SI for normal-hearing (NH) listeners based on a signal-to-noise ratio measure in the envelope domain (SNRenv), in the framework of the speech-based envelope power spectrum model (sEPSM, [20, 21]). These models have shown good agreement with measured data under a broad range of conditions, including stationary and modulated interferers, reverberation, and spectral subtraction. Despite the advances in modeling intelligibility in NH listeners, a broadly applicable model that can predict SI in hearing-impaired (HI) listeners is not yet available. As a first step towards such a model, this study investigates to what extent effects of hearing impairment on SI can be modeled in the sEPSM framework. Preliminary results show that, by only modeling the loss of audibility, the model cannot account for the higher speech reception threshold (SRT) of HI people in stationary noise compared to NH. However, this approach can, to some extent, account for the reduced ability of HI people to listen in the dips as measured by a reduced masking release (MR), where MR is denoted as the SRT benefit listeners obtain in fluctuating noise compared to stationary noise. The remaining causes of reduced MR that is not accounted for by the model could be due to additional effects of hearing impairment, such as broader auditory filters or deficits in temporal fine structure.

Native and non-native sentence comprehension in the presence of a competing talker

Objective correlates of pitch salience using pupillometry

Although objective correlates of pitch salience have been investigated in several neuroimaging studies, the results remain controversial. In the present study, a novel approach to objectively estimate pitch salience was used. Pupil dilation was measured as an indicator of the required effort in performing a pitch discrimination task for complex tones of varying pitch salience. It has been shown that cognitive processing demands of the task can be reflected in the pupil response, whereby pupil size dilates when cognitive load increases. The hypothesis was that pupil size would increase with increasing effort in performing the task and thus with decreasing pitch salience. A group of normal-hearing listeners first performed a behavioral pitch-discrimination experiment, where fundamental frequency difference limens (F₀ DLs) were measured as a function of F₀. Results showed that pitch salience of complex tones filtered in a high spectral region (1.5-3.5 kHz) increased with F₀. In a second experiment, listeners were presented with trials containing two reference complex tones with a fixed F₀ and a deviant tone with a larger F₀. Six conditions with different salience, defined by both the frequency region and F₀, were considered. Pupil size was measured for each condition, while the subjects’ task was to detect the deviants by pressing a response button. The expected trend was that pupil size would increase with decreasing salience. Results for musically trained listeners showed the expected trend, whereby pupil size significantly increased with
decreasing salience of the stimuli. Non-musically trained listeners showed, however, a smaller pupil size for the least salient condition as compared to a medium salient condition, probably due to a too demanding task.

Perception of a Sung Vowel as a Function of Frequency-Modulation Rate and Excursion in Normal-Hearing and Hearing-Impaired Listeners

Purpose: Frequency fluctuations in human voices can usually be described as coherent frequency modulation (FM). As listeners with hearing impairment (HI listeners) are typically less sensitive to FM than listeners with normal hearing (NH listeners), this study investigated whether hearing loss affects the perception of a sung vowel based on FM cues. Method: Vibrato maps were obtained in 14 NH and 12 HI listeners with different degrees of musical experience. The FM rate and FM excursion of a synthesized vowel, to which coherent FM was applied, were adjusted until a singing voice emerged. Results: In NH listeners, adding FM to the steady vowel components produced perception of a singing voice for FM rates between 4.1 and 7.5 Hz and FM excursions between 17 and 83 cents on average. In contrast, HI listeners showed substantially broader vibrato maps. Individual differences in map boundaries were, overall, not correlated with audibility or frequency selectivity at the vowel fundamental frequency, with no clear effect of musical experience. Conclusion: Overall, it was shown that hearing loss affects the perception of a sung vowel based on FM-rate and FM-excursion cues, possibly due to deficits in FM detection or discrimination or to a degraded ability to follow the rate of frequency changes.
Perceptual Effects of Dynamic Range Compression in Popular Music Recordings

There is a widespread belief that the increasing use of dynamic range compression in music mastering (the loudnesswar) deteriorates sound quality but experimental evidence of perceptual effects is lacking. In this study, normal hearing listeners were asked to evaluate popular music recordings in original versions and in remastered versions with higher levels of dynamic range compression. Surprisingly, we found no evidence of preference for the less compressed music. We also failed to find differences in ratings of perceived “depth” between the original and more compressed audio. A low degree of response consistency between different presentations of the same music suggests that listeners are less sensitive to even high levels of dynamic range compression than often argued.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Aalborg University
Authors: Hjortkjær, J. (Intern), Walther-Hansen, M. (Ekstern)
Number of pages: 5
Pages: 37-41
Publication date: 2014
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of Audio Engineering Society
Volume: 62
Issue number: 1/2
ISSN (Print): 1549-4950
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Web of Science (2017): Indexed Yes
Pitch coding of complex tones in the normal and impaired auditory system

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Bianchi, F. (Intern), Santurette, S. (Intern), Dau, T. (Intern)
Number of pages: 1
Publication date: 2014
Event: Poster session presented at International Hearing Aid Research Conference 2014, Lake Tahoe, United States.
Main Research Area: Technical/natural sciences
Electronic versions:
Bianchi_et_al_IHCON2014.pdf
Source: PublicationPreSubmission
Source-ID: 93687842
Publication: Research - peer-review › Journal article – Annual report year: 2014
Predicting speech release from masking through spatial separation in distance
Speech intelligibility models typically consist of a preprocessing part that transforms stimuli into some internal (auditory) representation and a decision metric that relates the internal representation to speech intelligibility. This study investigated speech intelligibility in conditions of spatial release from masking (SRM) where the masker is moved, on-axis, away from the target. Two binaural models, which use the conventional audio signal-to-noise ratio (SNR) in the decision metric, and two monaural models, using a decision metric based on the SNR in the envelope domain (SNRenv), were considered. The predictions were compared to data from Westermann et al. (2013, POMA, 19, 050156) in conditions where the target was located 0.5 m in front of the listener and the masker was presented at a distance of 0.5, 2, 5 or 10 m in front of the listener. The data showed an SRM of 10 dB when moving the masker from a distance of 0.5 m to a distance of 10 m. The long-term monaural model based on the SNRenv metric was able to account for most of the SRM data, whereas the models that used the audio SNR did not predict any SRM, even when they included an equalization-cancellation-like process. The short-term monaural model based on the SNRenv metric predicted a small SRM only in the noise-masker condition. The results suggest that true binaural processing is not always crucial to account for speech intelligibility in spatial conditions and that an SNR metric in the envelope domain appears to be more appropriate in conditions of on-axis spatial speech segregation than the conventional SNR. Additionally, none of the models considered grouping cues, which seem to play an important role in the conditions studied.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Chabot-Leclerc, A. (Intern), Dau, T. (Intern)
Number of pages: 6
Publication date: 2014

Refining a model of hearing impairment using speech psychophysics
The premise of this study is that models of hearing, in general, and of individual hearing impairment, in particular, can be improved by using speech test results as an integral part of the modeling process. A conceptual iterative procedure is presented which, for an individual, considers measures of sensitivity, cochlear compression, and phonetic confusions using the Diagnostic Rhyme Test (DRT) framework. The suggested approach is exemplified by presenting data from three hearing-impaired listeners and results obtained with models of the hearing impairment of the individuals. The work reveals that the DRT data provide valuable information of the damaged periphery and that the non-speech and speech data are complementary in obtaining the best model for an individual.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Boston University
Authors: Jepsen, M. L. (Intern), Dau, T. (Intern), Ghitza, O. (Ekstern)
Pages: EL179-EL85
Publication date: 2014
Main Research Area: Technical/natural sciences
Requirements for the evaluation of computational speech segregation systems

Recent studies on computational speech segregation reported improved speech intelligibility in noise when estimating and applying an ideal binary mask with supervised learning algorithms. However, an important requirement for such systems in technical applications is their robustness to acoustic conditions not considered during training. This study demonstrates that the spectro-temporal noise variations that occur during training and testing determine the achievable segregation performance. In particular, such variations strongly affect the identification of acoustical features in the system associated with perceptual attributes in speech segregation. The results could help establish a framework for a systematic evaluation of future segregation systems.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: May, T. (Intern), Dau, T. (Intern)
Number of pages: 7
Pages: 398-404
Publication date: 2014
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of the Acoustical Society of America
Volume: 136
Issue number: 6
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
Spontaneous otoacoustic emissions are generated by active oscillators clustered in frequency plateaus

It is commonly assumed that the active process linked to hair-cell motility is an important factor contributing to SOAEs. A chain of coupled, active and nonlinear oscillators with tonotopic organization can be used to account for key aspects of cochlear processing, including SOAEs and related phenomena when random irregularities of the mechanical parameters (roughness) are introduced. It was hypothesized that this roughness leads to sudden impedance mismatches leading to multiple reflections of the travelling wave in the cochlea. Recently it was shown [Wit & van Dijk, 2012; J. Acoust. Soc. Am. 132, 918–926] that a linear array of active oscillators with nearest neighbour coupling produces clusters of oscillators with a common oscillation frequency (frequency plateaus) and a preferred frequency separation. The frequency plateaus can also be entrained to the frequency of an external tone. Both of these aspects are properties found in SOAEs. In the present study it is investigated if frequency plateaus are also found in a TM which is able to generate realistic SOAEs and if these frequency plateaus can be used to explain the formation of SOAEs.

Temporal control and compensation for perturbed voicing feedback

Previous research employing a real-time auditory perturbation paradigm has shown that talkers monitor their own speech attributes such as fundamental frequency, vowel intensity, vowel formants, and fricative noise as part of speech motor control. In the case of vowel formants or fricative noise, what was manipulated is spectral information about the filter function of the vocal tract. However, segments can be contrasted by parameters other than spectral configuration. It is possible that the feedback system monitors phonation timing in the way it does spectral information. This study examined whether talkers exhibit a compensatory behavior when manipulating information about voicing. When talkers received feedback of the cognate of the intended voicing category (saying “tipper” while hearing “dipper” or vice versa), they changed the voice onset time and in some cases the following vowel.
The effect of interaural-time-difference fluctuations on apparent source width

For the perception of spaciousness, the temporal fluctuations of the interaural time differences (ITDs) and interaural level differences (ILDs) provide important binaural cues. One major characteristic of spatial perception is apparent source width (ASW), which describes the perceived width of a sound image. The temporal fluctuations of the binaural cues cause the signals at a listeners' ears to be decorrelated. Therefore, ASW has traditionally been measured by using the interaural cross-correlation (IACC). In particular, ITD fluctuations (below 2kHz) have been suggested to be the dominant cue for the perception of ASW. However, the contribution of the ITD statistics on the percept of ASW has not yet been clarified. In the present study, the impact of ITD fluctuations in different frequency bands on the perceived ASW was investigated. In a
psychoacoustic evaluation, a source signal was convolved with individual binaural room impulse responses (BRIRs) and presented to the listener via headphones. The obtained signals were passed through a gammatone filterbank with an analysis and synthesis stage which enabled the modification of the ITD fluctuation statistics in individual frequency bands. The ITD fluctuations of broadband noise stimuli were compressed while the effect of this compression on the ILD statistics was kept minimal. The IACC was kept the same for stimuli with compression below 2kHz and for the uncompressed noise which should lead to the same ASW percept in the two conditions. However, the psychoacoustic data showed a reduced ASW for the modified signals, particularly in conditions with an applied compression around 1 kHz. In contrast, above 2kHz, the compression had no effect on ASW, whereas the IACC increased. The results suggest that the broadband IACC can be a misleading objective measure of ASW and that ITD fluctuations around 1kHz are crucial for ASW perception.

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Acoustic Technology, Technical University of Denmark, Samsung Advanced Institute of Technology
Authors: Käsbach, J. (Intern), May, T. (Intern), Oskarsdottir, G. (Ekstern), Jeong, C. (Intern), Chang, J. (Ekstern)
Number of pages: 5
Publication date: 2014

**Host publication information**

Title of host publication: Proceedings of Forum Acusticum
BFI conference series: Forum Acusticum (5010988)
Main Research Area: Technical/natural sciences
Source: PublicationPreSubmission
Source-ID: 100044424
Publication: Research - peer-review › Article in proceedings – Annual report year: 2014

**The importance of binaural cues for the perception of apparent source width at different sound pressure levels**

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Käsbach, J. (Intern), May, T. (Intern), Le Goff, N. (Intern), Dau, T. (Intern)
Number of pages: 2
Publication date: 2014

**Host publication information**

Title of host publication: Proceedings of DAGA 2014
Main Research Area: Technical/natural sciences
Conference: DAGA 2014, Oldenburg, Germany, 10/03/2014 - 10/03/2014
Source: PublicationPreSubmission
Source-ID: 93780207
Publication: Research - peer-review › Article in proceedings – Annual report year: 2014

**The musical brain**

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Hjortkjær, J. (Intern)
Pages: 223-255
Publication date: 2014

**Host publication information**

Title of host publication: An Introduction to Neuroaesthetics. : The Neuroscientific Approach to Aesthetic Experience, Artistic Creativity, and Arts Appreciation
Publisher: Museum Tusculanums Forlag
Editor: Lauring, J. O.
ISBN (Print): 978-87-635-4140-4
Chapter: 7
Main Research Area: Technical/natural sciences
Source: PublicationPreSubmission
Source-ID: 110976158
The role of auditory spectro-temporal modulation filtering and the decision metric for speech intelligibility prediction

Speech intelligibility models typically consist of a preprocessing part that transforms stimuli into some internal (auditory) representation and a decision metric that relates the internal representation to speech intelligibility. The present study analyzed the role of modulation filtering in the preprocessing of different speech intelligibility models by comparing predictions from models that either assume a spectro-temporal (i.e., two-dimensional) or a temporal-only (i.e., one-dimensional) modulation filterbank. Furthermore, the role of the decision metric for speech intelligibility was investigated by comparing predictions from models based on the signal-to-noise envelope power ratio, SNRenv, and the modulation transfer function, MTF. The models were evaluated in conditions of noisy speech (1) subjected to reverberation, (2) distorted by phase jitter, or (3) processed by noise reduction via spectral subtraction. The results suggested that a decision metric based on the SNRenv may provide a more general basis for predicting speech intelligibility than a metric based on the MTF. Moreover, the one-dimensional modulation filtering process was found to be sufficient to account for the data when combined with a measure of across (audio) frequency variability at the output of the auditory preprocessing. A complex spectro-temporal modulation filterbank might therefore not be required for speech intelligibility prediction.
The role of temporal coherence in auditory stream segregation

The ability to perceptually segregate concurrent sound sources and focus one’s attention on a single source at a time is essential for the ability to use acoustic information. While perceptual experiments have determined a range of acoustic cues that help facilitate auditory stream segregation, it is not clear how the auditory system realizes the task. This thesis presents a study of the mechanisms involved in auditory stream segregation. Through a combination of psychoacoustic experiments, designed to characterize the influence of acoustic cues on auditory stream formation, and computational models of auditory processing, the role of auditory preprocessing and temporal coherence in auditory stream formation was evaluated. The computational model presented in this study assumes that auditory stream segregation occurs when sounds stimulate non-overlapping neural populations in a temporally incoherent manner. In the presented model, a physiologically inspired model of auditory preprocessing and perception was used to transform a sound signal into an auditory representation, and a subsequent temporal coherence analysis grouped frequency channels of the model together if they were stimulated in a temporally coherent manner. Based on this framework, the model was able to quantitatively predict perceptual experiments on stream segregation based on frequency separation and tone repetition rate, and onset and offset synchrony. Through the model framework, the influence of various processing stages on the stream segregation process was analysed. The model analysis showed that auditory frequency selectivity and physiological forward masking play a significant role in stream segregation based on frequency separation and tone rate. Secondly, the model analysis suggested that neural adaptation, and the resulting enhancement of neural responses to onsets, increases the sensitivity to onset synchrony for auditory stream formation. The effect of sound intensity on auditory stream formation was investigated, under the assumption that the wider auditory filters at high sound pressure levels should lead to a decreased ability to perceptually segregate sounds presented at high intensities. The results of listening experiments confirmed this hypothesis, showing that the minimum frequency separation required for stream segregation increases with increases in sound intensity. The computational model results also showed an increased tendency to group sounds presented at high intensities, but the size of the effect was overestimated relative to the experimental data, suggesting that the computational model does not fully reflect the auditory stream formation process. Lastly, an experimental paradigm designed to measure perceptual organization through an indirect, performance-based measure was investigated. This measure used comodulation masking release (CMR) to assess the conditions under which a loss of temporal coherence across frequency can lead to auditory stream segregation. The study indicated that CMR may be used as an indirect measure of stream segregation, and further supports the hypothesis that temporal coherence acts as a strong grouping cue. Overall, the findings of this thesis suggest that temporal coherence plays a significant role in the grouping of sounds.
into a single stream, and more generally, that a temporal coherence analysis may provide the framework for determining the perceptual organization of sounds into streams.

The shape of sounds: Audiovisual integration of visual shapes and musical sounds in the human brain

A multi-resolution envelope-power based model for speech intelligibility
The speech-based envelope power spectrum model (sEPSM) presented by Jørgensen and Dau [(2011). J. Acoust. Soc. Am. 130, 1475-1487] estimates the envelope power signal-to-noise ratio (SNRenv) after modulation-frequency selective processing. Changes in this metric were shown to account well for changes of speech intelligibility for normal-hearing listeners in conditions with additive stationary noise, reverberation, and nonlinear processing with spectral subtraction. In the latter condition, the standardized speech transmission index [(2003). IEC 60268-16] fails. However, the sEPSM is limited to conditions with stationary interferers, due to the long-term integration of the envelope power, and cannot account for increased intelligibility typically obtained with fluctuating maskers. Here, a multi-resolution version of the sEPSM is presented where the SNRenv is estimated in temporal segments with a modulation-filter dependent duration. The multi-resolution sEPSM is demonstrated to account for intelligibility obtained in conditions with stationary and fluctuating interferers, and noisy speech distorted by reverberation or spectral subtraction. The results support the hypothesis that the SNRenv is a powerful objective metric for speech intelligibility prediction. © 2013 Acoustical Society of America.
Ratings:

BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 0.848 SNIP 1.633
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 0.865 SNIP 1.647
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 0.752 SNIP 1.559
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 0.954 SNIP 1.749
Web of Science (2005): Indexed yes
Scopus rating (2004): SJR 0.77 SNIP 1.787
Web of Science (2004): Indexed yes
Scopus rating (2003): SJR 0.882 SNIP 1.712
Web of Science (2003): Indexed yes
Scopus rating (2002): SJR 0.87 SNIP 1.501
Web of Science (2002): Indexed yes
Scopus rating (2001): SJR 0.719 SNIP 1.467
Web of Science (2001): Indexed yes
Scopus rating (2000): SJR 0.621 SNIP 1.411
Web of Science (2000): Indexed yes
Scopus rating (1999): SJR 0.591 SNIP 1.319

Original language: English

Audition, Modulation, Reverberation, Signal to noise ratio, Speech intelligibility, Acoustic noise, Hearing
Analysis of the Auditory System via Sound Texture Synthesis

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: McWalter, R. I. (Intern), Dau, T. (Intern)
Pages: 1114-1117
Publication date: 2013

Host publication information
Title of host publication: Proceedings of the International Conference on Acoustics - AIA-DAGA 2013
Main Research Area: Technical/natural sciences
Conference: AIA-DAGA 2013 Conference on Acoustics, Merano, Italy, 18/05/2013 - 18/05/2013
Electronic versions:
2012_AiaDaga.pdf
Source: dtu
Source-ID: u::8249
Publication: Research - peer-review › Article in proceedings – Annual report year: 2013

Recent studies have attempted to control sound fields, and also to reduce room reflections with a circular or spherical array of loudspeakers. One of the attempts was to suppress sound waves propagating to the walls outside the array with a circular double-layer array of loudspeakers. The double-layer array represents a set of a monopole and a dipole in the Kirchhoff-Helmholtz integral equation, and thus the distance between these layers should be short compared with the wavelength. In practice, however, this condition is occasionally hard to satisfy because of the sizes of loudspeaker cabinets. In order to solve this problem, this study aims to examine several arrangements of a pair of loudspeakers that has a short distance between the acoustic centres. The effect of diffraction of sound waves due to the enclosure of another loudspeaker is investigated in simulations in terms of the position of the acoustic centre. As a result, it is shown that a loudspeaker has an approximately omni-directional radiation pattern at low frequencies in spite of the other loudspeaker cabinet, but the acoustic centre is shifted to the opposite direction of the cabinets.

General information
State: Published
Authors: Chang, J. (Intern), Agerkvist, F. T. (Intern), Olsen, M. (Ekstern)
Pages: 258-261
Publication date: 2013

Host publication information
Title of host publication: Proceedings of AES 52nd International Conference: Sound Field Control - Engineering and Perception
Publisher: Audio Engineering Society
Article number: 6-2
Main Research Area: Technical/natural sciences
Conference: 52nd International Conference: Sound Field Control, Guildford, United Kingdom, 02/09/2013 - 02/09/2013
Electronic versions:

Binaural dereverberation based on interaural coherence histograms
A binaural dereverberation algorithm is presented that utilizes the properties of the interaural coherence (IC) inspired by the concepts introduced in Allen et al. [J. Acoust. Soc. Am. 62, 912-915 (1977)]. The algorithm introduces a non-linear sigmoidal coherence-to-gain mapping that is controlled by an online estimate of the present coherence statistics. The algorithm automatically adapts to a given acoustic environment and provides a stronger dereverberation effect than the original method presented in Allen et al. [J. Acoust. Soc. Am. 62, 912-915 (1977)] in most acoustic conditions. The
The performance of the proposed algorithm was objectively and subjectively evaluated in terms of its impacts on the amount of reverberation and overall quality. A binaural spectral subtraction method based on Lebart et al. [Acta Acust. Acust. 87, 359-366 (2001)] and a binaural version of the original method of Allen et al. were considered as reference systems. The results revealed that the proposed coherence-based approach is most successful in acoustic scenarios that exhibit a significant spread in the coherence distribution where direct sound and reverberation can be segregated. This dereverberation algorithm is thus particularly useful in large rooms for short source-receiver distances.
Can comodulation masking release occur when frequency changes could promote perceptual segregation of the on-frequency and flanking bands?

A common characteristic of natural sounds is that the level fluctuations in different frequency regions are coherent. The ability of the auditory system to use this comodulation is shown when a sinusoidal signal is masked by a masker centred at the signal frequency (on-frequency masker, OFM) and one or more off-frequency components, commonly referred to as flanking bands (FBs). In general, the threshold of the signal masked by comodulated masker components is lower than when masked by masker components with uncorrelated envelopes or in the presence of the OFM only. This effect is commonly referred to as comodulation masking release (CMR). The present study investigates if CMR is also observed for a sinusoidal signal embedded in the OFM when the centre frequencies of the FBs are swept over time with a sweep rate of one octave per second. Both a common change of different frequencies and comodulation could serve as cues to indicate which of the stimulus components originate from one source. If the common fate of frequency components is the stronger binding cue, the sweeping FBs and the OFM with a fixed centre frequency should no longer form one auditory object and the CMR should be abolished. However, psychoacoustical results with normal-hearing listeners show that a CMR is also observed with sweeping components. The results are consistent with the hypothesis of wideband inhibition as the underlying physiological mechanism, as the CMR should only depend on the spectral position of the flanking bands relative to the inhibitory areas (as seen in physiological recordings using stationary flanking bands). Preliminary physiological results in the cochlear nucleus of the Guinea pig show that a correlate of CMR can also be found at this level of the auditory pathway with sweeping flanking bands.
Causal binary mask estimation for speech enhancement using sparsity constraints

While most single-channel noise reduction algorithms fail to improve speech intelligibility, the ideal binary mask (IBM) has demonstrated substantial intelligibility improvements for both normal- and impaired-hearing listeners. However, this approach exploits oracle knowledge of the target and interferer signals to preserve only the time-frequency regions that are target-dominated. Single-channel noise suppression algorithms trying to approximate the IBM using locally estimated signal-to-noise ratios without oracle knowledge have had limited success. Thought of in another way, the IBM exploits the disjoint placement of the target and interferer in time and frequency to create a time-frequency signal representation that is more sparse (i.e., has fewer non-zeros). In recent work (submitted to ICASSP 2013) we have introduced a novel time-frequency masking algorithm based on a sparse approximation algorithm from the signal processing literature. However, the algorithm employs a non-causal estimator. The present work introduces an improved de-noising algorithm that uses more realistic frame-based (causal) computations to estimate a binary mask.

Cochlear contributions to the precedence effect.

Normal-hearing individuals have sharply tuned auditory filters, and consequently their basilar-membrane (BM) impulse responses (IRs) have durations of several ms at frequencies in the range from 0 to 5 kHz. When presenting clicks that are several ms apart, the BM IRs to the individual clicks will overlap in time, giving rise to complex interactions that have not been fully understood in the human cochlea. The perceptual consequences of these BM IR interactions are of interest as lead-lag click pairs are often used to study localization and the precedence effect. The present study aimed at characterizing perceptual consequences of BM IR interactions in individual listeners based on click-evoked otoacoustic emissions (CEOAEs) and auditory brainstem responses (ABRs). Lag suppression, denoting the level difference between the CEOAE or wave-V response amplitude evoked by the first and the second clicks, was observed for inter-click intervals (ICIs) between 1 and 4 ms. Behavioral correlates of lag suppression were obtained for the same individuals by investigating the percept of the lead-lag click pairs presented either monaurally or binaurally. The click pairs were shown to give rise to fusion (i.e., the inability to hear out the second click in a lead-lag click pair), regardless of monaural or binaural presentation. In both cases, the ICI range where the percept was a fused image correlated well with the ICI range...
for which monaural lag suppression occurred in the CEOAE and ABR (i.e., for ICIs below 4.3 ms). Furthermore, the lag suppression observed in the wave-V amplitudes to binaural stimulation did not show additional contributions to the lag suppression obtained monaurally, suggesting that peripheral lag suppression up to the level of the brainstem is dominant in the perception of the precedence effect.

**General information**

**State:** Published  
**Organisations:** Department of Electrical Engineering, Hearing Systems, Boston University  
**Authors:** Verhulst, S. (Ekstern), Bianchi, F. (Intern), Dau, T. (Intern)  
**Pages:** 283-291  
**Publication date:** 2013  

**Host publication information**  
**Title of host publication:** Advances in Experimental Medicine and Biology  
**Volume:** 787  
**Publisher:** Springer Science+Business Media  
**ISBN (Print):** 978-1-4614-1589-3  
**ISBN (Electronic):** 978-1-4614-1590-9  
**Chapter:** 32  
**Series:** Advances in Experimental Medicine and Biology  
**ISSN:** 0065-2598  
**Main Research Area:** Technical/natural sciences  
**Otorhinolaryngology, Neurosciences, Animal Physiology, Biophysics and Biological Physics, Neurobiology**  
**DOIs:**  
10.1007/978-1-4614-1590-9_32  
Source: dtu  
Source-ID: n:oai:DTIC-ART:pubmed/387371342::31773  
**Publication:** Research - peer-review » Book chapter – Annual report year: 2013

**Contribution of envelope periodicity to release from speech-on-speech masking**  
Masking release (MR) is the improvement in speech intelligibility for a fluctuating interferer compared to stationary noise. Reduction in MR due to vocoder processing is usually linked to distortions in the temporal fine structure of the stimuli and a corresponding reduction in the fundamental frequency (F0) cues. However, it is unclear if envelope periodicity related to F0, produced by the interaction between unresolved harmonics, contributes to MR. In the present study, MR was determined from speech reception thresholds measured in the presence of stationary speech-shaped noise and a competing talker. Two types of processing were applied to the stimuli: (1) An amplitude- and frequency-modulated vocoder attenuated the envelope periodicity and (2) high-pass (HP) filtering (cutoff=4500 Hz) reduced the influence of F0-related information from low-order resolved harmonics. When applied individually, MR was unaffected by HP filtering, but slightly reduced when envelope periodicity was attenuated. When both were applied, MR was strongly reduced. Thus, the results indicate that F0-related information is crucial for MR, but that it is less important whether the F0-related information is conveyed by low-order resolved harmonics or by envelope periodicity as a result of unresolved harmonics. Further, envelope periodicity contributes substantially to MR.

**General information**

**State:** Published  
**Organisations:** Department of Electrical Engineering, Hearing Systems, Technical University of Denmark  
**Authors:** Christiansen, C. (Ekstern), MacDonald, E. (Intern), Dau, T. (Intern)  
**Pages:** 2197–2204  
**Publication date:** 2013  
**Main Research Area:** Technical/natural sciences

**Publication information**  
**Journal:** Journal of the Acoustical Society of America  
**Volume:** 134  
**Issue number:** 3  
**ISSN (Print):** 0001-4966  
**Ratings:**  
BFI (2018): BFI-level 2  
Web of Science (2018): Indexed yes  
BFI (2017): BFI-level 2  
Web of Science (2017): Indexed yes  
BFI (2016): BFI-level 2  
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Do perceptual consequences of spontaneous otoacoustic emissions reflect a central plasticity effect?
Frequency difference limens (FDLs) have been found to improve in the vicinity of spontaneous otoacoustic emissions (SOAEs). This effect has been observed ipsilaterally and contralaterally to the emission ear, suggesting that prolonged...
ongoing stimulation of nerve cells tuned to the SOAE frequency lead to a central oversensitivity to that frequency (Norena et al., 2002, Hear. Res.). However, it is known that a tone close to an SOAE frequency can "entrain" the emission to oscillate at the tone frequency (Long and Tubis, 1988, Hear. Res.), thus FDLs near SOAEs might also be affected by this peripheral process. An alternative hypothesis explaining FDL performance in terms of peripheral entrainment and beating between external tones and SOAEs is proposed here. SOAE entrainment patterns were obtained in seven subjects with an ipsilateral SOAE and no neighboring contralateral SOAE. Ipsilateral FDLs were measured at frequencies covering the individual entrainment and beating regions. Ipsilateral FDLs were lowest in the entrainment region and worsened significantly when beating occurred. However, no changes in contralateral FDLs were found. Contralateral FDLs also remained unaffected by continuous ipsilateral presentation of a pure tone aimed at emulating an SOAE. These findings suggest a peripheral rather than central plasticity origin for perceptual consequences of SOAEs.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Technical University of Denmark, Boston University
Authors: Hansen, R. (Ekstern), Santurette, S. (Intern), Verhulst, S. (Ekstern)
Publication date: 2013
Event: Poster session presented at International Symposium on Auditory and Audiological Research, Nyborg, Denmark.
Main Research Area: Technical/natural sciences
Electronic versions:
Hansen_ISAAR_2013_poster.pdf
Source: dtu
Source-ID: u::9004
Publication: Research - peer-review › Poster – Annual report year: 2013

Effects of spontaneous otoacoustic emissions on frequency discrimination
When an external tone is presented in proximity to the frequency of a spontaneous otoacoustic emission (SOAE), the SOAE typically synchronizes to the external tone, a phenomenon known as "entrainment". As the tone moves further away from the SOAE frequency, beating patterns between the SOAE and the pure tone occur (Long, Hear. Res. 119, 1998). This study investigated perceptual consequences of SOAE beating and entrainment on the frequency difference limen (FDL), which has been found to improve near an SOAE. SOAE entrainment patterns were obtained for six subjects with a strong SOAE in the ipsilateral ear and no SOAE in the corresponding frequency range of the contralateral ear. Hearing thresholds and FDLs were measured ipsi- and contralaterally for nine frequencies covering the entrainment and beating regions of the SOAE. FDLs systematically improved in the entrainment region, worsened when beating occurred, and improved again for frequencies further away from the SOAE. No improvement in FDL was found in any of the contralateral ears tested, suggesting that the effect is of peripheral, rather than of central, origin. The results contradict an earlier hypothesis stating that FDL performance near SOAE frequencies is governed by a central oversensitivity to the SOAE frequency.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Technical University of Denmark
Authors: Hansen, R. (Ekstern), Santurette, S. (Intern), Verhulst, S. (Intern)
Number of pages: 9
Publication date: 2013
Conference: 21st International Congress on Acoustics, Montreal, Canada, 02/06/2013 - 02/06/2013
Main Research Area: Technical/natural sciences
Publication information
Journal: Meetings on Acoustics. Proceedings
Volume: 19
ISSN (Print): 1939-800X
Ratings:
Scopus rating (2016): SNIP 0.124 CiteScore 0.15
Scopus rating (2015): SJR 0.218 SNIP 0.147
Scopus rating (2014): SJR 0.188 SNIP 0.105
Scopus rating (2013): SJR 0.138 SNIP 0.243
ISI indexed (2013): ISI indexed no
Scopus rating (2012): SJR 0.144 SNIP 0.21
ISI indexed (2012): ISI indexed no
Scopus rating (2011): SJR 0.141 SNIP 0.313
ISI indexed (2011): ISI indexed no
Scopus rating (2010): SJR 0.163 SNIP 0.09
Efficient estimates of cochlear hearing loss parameters in individual listeners

It has been suggested that the level corresponding to the knee-point of the basilar membrane (BM) input/output (I/O) function can be used to estimate the amount of inner- and outer hair-cell loss (IHL, OHL) in listeners with a moderate cochlear hearing impairment Plack et al. (2004). According to Jepsen and Dau (2011) IHL + OHL = HLT [dB], where HLT stands for total hearing loss. Hence having estimates of the total hearing loss and OHC loss, one can estimate the IHL. In the present study, results from forward masking experiments based on temporal masking curves (TMC; Nelson et al., 2001) are presented and used to estimate the knee-point level and the compression ratio of the I/O function. A time-efficient paradigm based on the single-interval-up-down method (SIUD; Lecluyse and Meddis (2009)) was used. In contrast with previous studies, the present study used only on-frequency TMCs to derive estimates of the knee-point level. Further, it is explored whether it is possible to estimate the compression ratio using only on-frequency TMCs. 10 normal-hearing and 10 hearing-impaired listeners (with mild-to-moderate sensorineural hearing loss) were tested at 1, 2 and 4 kHz. The results showed a reasonable reliability and may be applicable to individualized hearing-aid fitting. © 2013 Acoustical Society of America.
Environment-aware ideal binary mask estimation using monaural cues

We present a monaural approach to speech segregation that estimates the ideal binary mask (IBM) by combining amplitude modulation spectrogram (AMS) features, pitch-based features and speech presence probability (SPP) features derived from noise statistics. To maintain a high mask estimation accuracy in the presence of various background noises, the system employs environment-specific segregation models and automatically selects the appropriate model for a given input signal. Furthermore, instead of classifying each timefrequency (T-F) unit independently, the a posteriori probabilities of speech and noise presence are evaluated by considering adjacent TF units. The proposed system achieves high classification accuracy.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: May, T. (Intern), Dau, T. (Intern)
Number of pages: 4
Publication date: 2013

Host publication information
Title of host publication: 2013 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics
Publisher: IEEE
Main Research Area: Technical/natural sciences
Ideal binary mask estimation, Speech segregation, Background noise classification
DOIs: 10.1109/WASPAA.2013.6701821
Source: dtu
Source-ID: u::8219
Publication: Research - peer-review › Article in proceedings – Annual report year: 2013

Estimates of human cochlear tuning derived from DPOAE reflection-component delays

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Southern California, House Research Institute, Harvard Medical School
Authors: Abdala, C. (Ekstern), Guerit, F. M. L. P. (Intern), Luo, P. (Ekstern), Shera, C. A. (Ekstern)
Number of pages: 1
Publication date: 2013
Event: Poster session presented at 36th Annual Midwinter Meeting of the Association for Research in Otolaryngology, Baltimore, MD, United States.
Main Research Area: Technical/natural sciences
Electronic versions:
Abdala et al. - 2013 - ESTIMATES OF HUMAN COCHLEAR TUNING DERIVED FROM DPOAE REFLECTION-COMPONENT DELAYS.pdf

Bibliographical note
Conference poster (16 February 2013)
36th Annual Midwinter Meeting of the Association for Research in Otolaryngology
Source: dtu
Source-ID: u::9167
Publication: Research - peer-review › Poster – Annual report year: 2013

Experimental Evidence for a Cochlear Source of the Precedence Effect
The precedence effect (PE) refers to the dominance of directional information carried by a direct sound (lead) over the spatial information contained in its multiple reflections (lags) in sound localization. Although the processes underlying the PE have been largely investigated, the extent to which peripheral versus central auditory processes contribute to this perceptual phenomenon has remained unclear. The present study investigated the contribution of peripheral processing to the PE through a comparison of physiological and psychoacoustical data in the same human listeners. The psychoacoustical experiments, comprising a fusion task, an interaural time difference detection task and a lateralization task, demonstrated a time range from 1 to 4.6–5 ms, in which the PE operated (precedence window). Click-evoked otoacoustic emissions (CEOAEs) were recorded in both ears to investigate the lead–lag interactions at the level of the basilar membrane (BM) in the cochlea. The CEOAE-derived peripheral and monaural lag suppression was largest for ICIs of 1–4 ms. Auditory-evoked brainstem responses (ABRs) were used to investigate monaural and binaural lag suppression at the brainstem level. The responses to monaural stimulation reflected the peripheral lag suppression observed in the CEOAE results, while the binaural brainstem responses did not show any substantial contribution of binaural processes to
monaural lag suppression. The results demonstrated that the lag suppression occurring at the BM in a time range from 1 to 4 ms, as indicated by the suppression of the lag-CEOAE, was the source of the reduction in the lag-ABRs and a possible peripheral contributor to the PE for click stimuli.
Experimental validation of sound field control with a circular double-layer array of loudspeakers.

This paper is concerned with experimental validation of a recently proposed method of controlling sound fields with a circular double-layer array of loudspeakers [Chang and Jacobsen, J. Acoust. Soc. Am. 131(6), 4518-4525 (2012)]. The double-layer of loudspeakers is realized with 20 pairs of closed-box loudspeakers mounted back-to-back. Source strengths are obtained with several solution methods by modeling loudspeakers as a weighted combination of monopoles and dipoles. Sound pressure levels of the controlled sound fields are measured inside and outside the array in an anechoic room, and performance indices are calculated. The experimental results show that a method of combining pure contrast maximization with a pressure matching technique provides only a small error in the listening zone between the desired and the reproduced fields, and at the same time reduces the sound level in the quiet zone as expected in the simulation studies well above the spatial Nyquist frequency except at a few frequencies. It is also shown that errors in the positions of the loudspeakers can be critical to the results at frequencies where the distance between the inner and the outer array is close to half a wavelength.
Extracting Sound-Source-Distance Information from Binaural Signals

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Patras, Technical University of Denmark
Authors: Georganti, E. (Ekstern), May, T. (Intern), van de Par, S. (Ekstern), Mourjopoulos, J. (Ekstern)
Pages: 171-199
Publication date: 2013

**Host publication information**

Title of host publication: The Technology of Binaural Listening : Modern Acoustics and Signal Processing
Publisher: Springer
Editor: Blauert, J.
Main Research Area: Technical/natural sciences
Source: dtu
Source-ID: u::7678
Publication: Research - peer-review › Book chapter – Annual report year: 2013
Frequency-Modulation Vowel Maps in Normal-Hearing and Hearing-Impaired Listeners

**General information**
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Oticon A/S
Authors: Santurette, S. (Intern), Vatti, M. (Intern), Pontoppidan, N. H. (Ekstern), Dau, T. (Intern)
Number of pages: 1
Pages: 233-233
Publication date: 2013

**Host publication information**
Title of host publication: Abstract Book
Publisher: Association for Research in Otolaryngology

**Electronic versions:**
ARO2013FinalAbstractBook.pdf

Intelligibility of speech produced in temporally modulated noise

**General information**
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Jade Hochschule, Oldenburg
Authors: MacDonald, E. (Intern), Raufer, S. (Ekstern)
Pages: 1297-1300
Publication date: 2013

**Host publication information**
Title of host publication: Proceedings of the International Conference on Acoustics
Main Research Area: Technical/natural sciences
Conference: AIA-DAGA 2013 Conference on Acoustics, Merano, Italy, 18/05/2013 - 18/05/2013
Source: dtu
Source-ID: u::7863
Publication: Research - peer-review › Article in proceedings – Annual report year: 2013

Interaural bimodal pitch matching with two-formant vowels
For bimodal patients, with a hearing aid (HA) in one ear and a cochlear implant (CI) in the opposite ear, usually a default frequency-to-electrode map is used in the CI. This assumes that the human brain can adapt to interaural place-pitch
mismatches. This “one-size-fits-all” method might be partly responsible for the large variability of individual bimodal benefit. Therefore, knowledge about the location of the electrode array is an important prerequisite for optimum fitting. Theoretically, the electrode location can be determined from CT-scans. However, these are often not available in audiological practice. Behavioral pitch matching between the two ears has also been suggested, but has been shown to be tedious and unreliable. Here, an alternative method using two-formant vowels was developed and tested with a vocoder system simulating different CI insertion depths. The hypothesis was that patients may more easily identify vowels than perform a classical pitch-matching task. A spectral shift is inferred by comparing vowel spaces, measured by presenting the first formant in the HA and the second either in the HA or the CI. Preliminary results suggest that pitch mismatches can be derived from such vowel spaces. In order to take auditory adaptation in individual patients into account, the method will be tested with CI patients with contralateral residual hearing.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Advanced Bionics GmbH
Authors: Guerit, F. M. L. P. (Intern), Chalupper, J. (Ekstern), Santurette, S. (Intern), Arweiler, I. (Ekstern), Dau, T. (Intern)
Publication date: 2013
Event: Poster session presented at International Symposium on Auditory and Audiological Research, Nyborg, Denmark.
Main Research Area: Technical/natural sciences
Electronic versions:
Guerit_posterISAAR.pdf
Source: dtu
Source-ID: u::9003
Publication: Research - peer-review › Poster – Annual report year: 2013

Interaural bimodal pitch matching with two-formant vowels
For bimodal patients, with a hearing aid (HA) in one ear and a cochlear implant (CI) in the opposite ear, usually a default frequency-to-electrode map is used in the CI. This assumes that the human brain can adapt to interaural place-pitch mismatches. This “one-size-fits-all” method might be partly responsible for the large variability of individual bimodal benefit. Therefore, knowledge about the location of the electrode array is an important prerequisite for optimum fitting. Theoretically, the electrode location can be determined from CT-scans. However, these are often not available in audiological practice. Behavioral pitch matching between the two ears has also been suggested, but has been shown to be tedious and unreliable. Here, an alternative method using two-formant vowels was developed and tested with a vocoder system simulating different CI insertion depths. The hypothesis was that patients may more easily identify vowels than perform a classical pitch-matching task. A spectral shift is inferred by comparing vowel spaces, measured by presenting the first formant in the HA and the second either in the HA or the CI. Preliminary results suggest that pitch mismatches can be derived from such vowel spaces. In order to take auditory adaptation in individual patients into account, the method will be tested with CI patients with contralateral residual hearing.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Advanced Bionics GmbH
Authors: Guérit, F. (Intern), Chalupper, J. (Ekstern), Santurette, S. (Intern), Arweiler, I. (Ekstern), Dau, T. (Intern)
Number of pages: 8
Publication date: 2013

Host publication information
Title of host publication: Proceedings of ISAAR 2013 : Auditory Plasticity – Listening with the Brain
Editors: Dau, T., Santurette, S., Dalsgaard, J. C., Tranebjærg, L., Andersen, T., Poulsen, T.
ISBN (Print): 978 -8 7 -9 90013 -4-7
Main Research Area: Technical/natural sciences
Electronic versions:
333_340_Guerit.pdf
Publication: Research - peer-review › Article in proceedings – Annual report year: 2014

Level dependency of auditory stream segregation

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Jade Hochschule, Oldenburg
Authors: Hauth, C. (Ekstern), Christiansen, S. K. (Intern), Dau, T. (Intern)
Publication date: 2013

Host publication information
Title of host publication: Proceedings of the Deutsche Gesellschaft für Akustik - 2013
Masking Release for Sweeping Masker Components with Correlated Envelopes

To separate sounds from different sound sources, common properties of natural sounds are used by the auditory system, such as coherent temporal envelope fluctuations and correlated changes of frequency in different frequency regions. The present study investigates how the auditory system processes a combination of these cues using a generalized comodulation masking release (CMR) paradigm. CMR is the effect of a better signal detectability in the presence of comodulated maskers than in the presence of maskers with uncorrelated envelope fluctuations across frequencies. Using a flanking-band paradigm, the results of the first experiment of the present study show that CMR is still observed for the masker and the signal coherently sweeping up or down in frequency over time, up to a sweep rate of six octaves per second. Motivated by the successful modeling of CMR using filters sensitive to temporal modulations and recent physiological evidence of spectro-temporal modulation filters, the second experiment investigates whether CMR is also observed for spectro-temporal masker modulations generated using time-shifted versions of the masker envelope for each component. The thresholds increase as soon as the temporally coherent masker modulation is changed to a spectro-temporal masker modulation.
Maximum Acceptable Vibrato Excursion as a Function of Vibrato Rate in Musicians and Non-musicians

Human vibrato is mainly characterized by two parameters: vibrato extent and vibrato rate. These parameters have been found to exhibit an interaction both in physical recordings of singers’ voices and in listener’s preference ratings. This study was concerned with the way in which the maximum acceptable vibrato excursion varies as a function of vibrato rate in normal-hearing (NH) musicians and non-musicians. Eight NH musicians and six non-musicians adjusted the maximum vibrato excursion of a synthesized vowel for vibrato rates between 3 and 8 Hz. Individual thresholds varied across vibrato rate and, in most listeners, exhibited a peak at medium vibrato rates (5–7 Hz). Large across-subject variability was observed, and no significant effect of musical experience was found. Overall, most listeners were not solely sensitive to the vibrato excursion and there was a listener-dependent rate for which larger vibrato excursions were favored. The observed interaction between maximum excursion thresholds and vibrato rate may be due to the listeners’ judgments relying on cues provided by the rate of frequency changes (RFC) rather than excursion per se. Further studies are needed to evaluate the contribution of the RFC to vibrato perception and the possible effects of age and hearing impairment.

Modeling auditory evoked potentials to complex stimuli

The auditory evoked potential (AEP) is an electrical signal that can be recorded from electrodes attached to the scalp of a human subject when a sound is presented. The signal is considered to reflect neural activity in response to the acoustic stimulation and is a well established clinical and research tool to objectively assess the function and integrity of the auditory nervous system. However, the physiological generation of AEPs represents a complicated interaction between linear and nonlinear cochlear and neural processes and is not well understood in humans. This thesis presents and evaluates a phenomenological model of AEP generation that can predict key experimental observations of recorded AEPs. The purpose of the study was to investigate the role of the different stages of auditory signal processing and their effects on AEP generation.

In recent years, there has been a push both clinically and in research towards using realistic and complex stimuli, such as speech, to electrophysiologically assess the human hearing. However, to interpret the AEP generation to complex sounds, the potential patterns in response to simple stimuli needs to be understood. Therefore, the model was used to simulate
auditory brainstem responses (ABRs) evoked by classic stimuli like clicks, tone bursts and chirps. The ABRs to these simple stimuli were compared to literature data and the model was shown to predict the frequency dependence of tone-burst ABR wave-V latency and the level-dependence of ABR waveV amplitude for clicks and chirps varying sweeping rates. The model was also evaluated based on ABR recordings evoked by speech syllables, and was shown to account for the differences in the responses observed between the stimuli. It was demonstrated that the generation of the syllable-evoked ABRs was highly influenced by cochlear and afferent neural processing, which supported the importance of cochlear processing for the generation of AEPs.

A second major contribution of this study was the investigation of whether auditory steady-state responses (ASSRs) can be used to assess human cochlear compression. Sensorineural hearing impairments is commonly associated with a loss of outer hair-cell functionality, and a measurable consequence is the decreased amount of cochlear compression at frequencies corresponding to the damaged locations in the cochlea. In clinical diagnostics, a fast and objective measure of local cochlear compression would be of great benefit, as a more precise diagnose of the deficits underlying a potential hearing impairment in both infants and adults could be obtained. It was demonstrated in this thesis, via experimental recordings and supported by model simulations, that the growth of the ASSR amplitude with stimulus level can indeed be used as such an estimate of local cochlear compression.

Modeling Horizontal Localization of Complex Sounds in the Impaired and Aided Impaired Auditory System

The relative contributions of within-channel and across-channel processes to perceptual comodulation masking release (CMR) were investigated in the framework of an auditory processing model. A generalized version of the computational auditory signal processing and perception model [CASP; Jepsen et al., J. Acoust. Soc. Am. 124, 422-438 (2008)] was used and extended by an across-channel modulation processing stage according to Piechowiak et al. [J. Acoust. Soc. Am. 121, 2111-2126 (2007)]. Five experimental paradigms were considered: CMR with a broadband noise masker as a function of the masker spectrum level; CMR with four widely spaced flanking bands (FBs) varying in overall level; CMR with one FB varying in frequency and level relative to the on-frequency band (OFB); CMR with one FB varying in frequency; and CMR as a function of the number of FBs. The predictions suggest that at least three different mechanisms contribute to overall CMR in the considered conditions: (1) a within-channel process based on changes in the envelope characteristic due to the addition of the signal to the masker; (2) a within-channel process based on nonlinear peripheral processing of the OFB’s envelope caused by the FB(s); and (3) an across-channel process that is robust across
presentation levels but relatively small (2-5 dB).

**General information**
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Oldenburg
Authors: Dau, T. (Intern), Piechowiak, T. (Intern), Ewert, S. D. (Ekstern)
Pages: 350-364
Publication date: 2013
Main Research Area: Technical/natural sciences

**Publication information**
Volume: 133
Issue number: 1
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 0.848 SNIP 1.633
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 0.865 SNIP 1.647
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 0.752 SNIP 1.559
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 0.954 SNIP 1.749
Web of Science (2005): Indexed yes
Modelling speech intelligibility in adverse conditions.

Jørgensen and Dau (J Acoust Soc Am 130:1475-1487, 2011) proposed the speech-based envelope power spectrum model (sEPSM) in an attempt to overcome the limitations of the classical speech transmission index (STI) and speech intelligibility index (SII) in conditions with nonlinearly processed speech. Instead of considering the reduction of the temporal modulation energy as the intelligibility metric, as assumed in the STI, the sEPSM applies the signal-to-noise ratio in the envelope domain (SNRenv). This metric was shown to be the key for predicting the intelligibility of reverberant speech as well as noisy speech processed by spectral subtraction. The key role of the SNRenv metric is further supported here by the ability of a short-term version of the sEPSM to predict speech masking release for different speech materials and modulated interferers. However, the sEPSM cannot account for speech subjected to phase jitter, a condition in which the spectral structure of the intelligibility of speech signal is strongly affected, while the broadband temporal envelope is kept largely intact. In contrast, the effects of this distortion can be predicted successfully by the spectro-temporal modulation index (STMI) (Elhilali et al., Speech Commun 41:331-348, 2003), which assumes an explicit analysis of the spectral “ripple” structure of the speech signal. However, since the STMI applies the same decision metric as the STI, it fails to account for spectral subtraction. The results from this study suggest that the SNRenv might reflect a powerful decision metric, while some explicit across-frequency analysis seems crucial in some conditions. How such across-frequency analysis is “realized” in the auditory system remains unresolved.
Mismatch Negativity (MMN) is a component in the auditory Event-Related Potential (ERP) that is elicited by a change in the auditory percept. It has been shown that the McGurk illusion can induce a MMN. We conducted an experiment in which the MMN could be elicited by the McGurk illusion induced by visual speech with either upright (unaltered) or vertically reversed mouth area. In a preliminary analysis, we found a Mismatch Negativity component induced by the McGurk illusion for 6 of 17 participants at electrode Cz when the mouth area was upright. In comparison, these participants produced no Mismatch Negativity when the mouth was reversed. These findings mirrored behavioral findings in the same subjects of a strong McGurk response for normal audiovisual speech, which was greatly reduced for stimuli with reversed mouth area.

**General information**

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems, Hearing Systems, Centre for Applied Hearing Research, Department of Electrical Engineering
Authors: Eskelund, K. (Intern), Andersen, T. (Intern)
Pages: 133–134
Publication date: 2013
Conference: 14th International Multisensory Research Forum (IMRF 2013), Jerusalem, Israel, 03/06/2013 - 03/06/2013
Main Research Area: Technical/natural sciences

**Publication information**

Journal: Multisensory Research
Volume: 26
Issue number: 0
ISSN (Print): 2213-4794
Ratings:
Web of Science (2018): Indexed yes
Web of Science (2017): Indexed Yes
Scopus rating (2016): SJR 0.611 SNIP 0.544 CiteScore 1.55
Scopus rating (2015): SJR 0.672 SNIP 0.549 CiteScore 1.22
Scopus rating (2014): SJR 0.655 SNIP 0.837 CiteScore 0.76
Scopus rating (2013): SJR 0.917 SNIP 0.675
Web of Science (2013): Indexed yes
Scopus rating (2012): SJR 0.807 SNIP 0.702
Scopus rating (2011): SJR 0.699 SNIP 0.944
Scopus rating (2010): SJR 0.664 SNIP 1.006
Scopus rating (2009): SJR 0.583 SNIP 0.955
Scopus rating (2008): SJR 0.782 SNIP 0.969
Scopus rating (2007): SJR 0.556 SNIP 0.645
Scopus rating (2006): SJR 0.764 SNIP 0.794
Scopus rating (2005): SJR 0.651 SNIP 0.613
Scopus rating (2004): SJR 0.633 SNIP 0.544
Scopus rating (2003): SJR 0.814 SNIP 0.767
Scopus rating (2002): SJR 0.841 SNIP 0.688
Scopus rating (2001): SJR 0.841 SNIP 0.539
Scopus rating (2000): SJR 1.111 SNIP 0.912
Scopus rating (1999): SJR 0.944 SNIP 0.929
Original language: English
Audiovisual integration, Speech perception, McGurk illusion, Event-related potentials, Mismatch negativity
Source: dtu
Source-ID: u::7612
Publication: Research - peer-review › Conference abstract in journal – Annual report year: 2013

**Multimodal speech perception**

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Queen's University
Authors: Alsius, A. (Ekstern), MacDonald, E. (Intern), Munhall, K. (Ekstern)
Number of pages: 77
Publication date: 2013
Multivoxel Patterns Reveal Functionally Differentiated Networks Underlying Auditory Feedback Processing of Speech

The everyday act of speaking involves the complex processes of speech motor control. An important component of control is monitoring, detection, and processing of errors when auditory feedback does not correspond to the intended motor gesture. Here we show, using fMRI and converging operations within a multivoxel pattern analysis framework, that this sensorimotor process is supported by functionally differentiated brain networks. During scanning, a real-time speech-tracking system was used to deliver two acoustically different types of distorted auditory feedback or unaltered feedback while human participants were vocalizing monosyllabic words, and to present the same auditory stimuli while participants were passively listening. Whole-brain analysis of neural-pattern similarity revealed three functional networks that were differentially sensitive to distorted auditory feedback during vocalization, compared with during passive listening. One network of regions appears to encode an “error signal” regardless of acoustic features of the error: this network, including right angular gyrus, right supplementary motor area, and bilateral cerebellum, yielded consistent neural patterns across acoustically different, distorted feedback types, only during articulation (not during passive listening). In contrast, a frontotemporal network appears sensitive to the speech features of auditory stimuli during passive listening; this preference for speech features was diminished when the same stimuli were presented as auditory concomitants of vocalization. A third network, showing a distinct functional pattern from the other two, appears to capture aspects of both neural response profiles. Together, our findings suggest that auditory feedback processing during speech motor control may rely on multiple, interactive, functionally differentiated neural systems.
Neural correlates of pitch salience using fMRI

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Copenhagen University Hospital
Authors: Bianchi, F. (Intern), Santurette, S. (Intern), Dau, T. (Intern), Hjortkjaer, J. (Intern), Siebner, H. (Ekstern)
Publication date: 2013
Event: Poster session presented at 7th Arches meeting, Paris, France.
Main Research Area: Technical/natural sciences
Electronic versions:
abstract_poster_ARCHES.pdf
Source-ID: u::9504
Publication: Research - peer-review › Poster – Annual report year: 2013

Objective measures of binaural masking level differences and comodulation masking release based on late auditory evoked potentials
The audibility of important sounds is often hampered due to the presence of other masking sounds. The present study investigates if a correlate of the audibility of a tone masked by noise is found in late auditory evoked potentials measured from human listeners. The audibility of the target sound at a fixed physical intensity is varied by introducing auditory cues of (i) interaural target signal phase disparity and (ii) coherent masker level fluctuations in different frequency regions. In agreement with previous studies, psychoacoustical experiments showed that both stimulus manipulations result in a masking release (i:...
binaural masking level difference; ii: comodulation masking release) compared to a condition where those cues are not present. Late auditory evoked potentials (N1, P2) were recorded for the stimuli at a constant masker level, but different signal levels within the same set of listeners who participated in the psychoacoustical experiment. The data indicate differences in N1 and P2 between stimuli with and without interaural phase disparities. However, differences for stimuli with and without coherent masker modulation were only found for P2, i.e., only P2 is sensitive to the increase in audibility, irrespective of the cue that caused the masking release. The amplitude of P2 is consistent with the psychoacoustical finding of an addition of the masking releases when both cues are present. Even though it cannot be concluded where along the auditory pathway the audibility is represented, the P2 component of auditory evoked potentials is a candidate for an objective measure of audibility in the human auditory system.

**General information**
- **State:** Published
- **Organisations:** Department of Electrical Engineering, Hearing Systems, University College London, Otto-von-Guericke University Magdeburg
- **Authors:** Epp, B. (Intern), Yasin, I. (Ekstern), Verhey, J. L. (Ekstern)
- **Pages:** 21-28
- **Publication date:** 2013
- **Main Research Area:** Technical/natural sciences

**Publication information**
- **Journal:** Hearing Research
- **Volume:** 306
- **ISSN (Print):** 0378-5955
- **Ratings:**
  - BFI (2018): BFI-level 2
  - Web of Science (2018): Indexed yes
  - BFI (2017): BFI-level 2
  - Web of Science (2017): Indexed Yes
  - BFI (2016): BFI-level 2
  - Web of Science (2016): Indexed yes
  - BFI (2015): BFI-level 1
  - Scopus rating (2015): SJR 1.857 SNIP 1.478 CiteScore 3.28
  - BFI (2014): BFI-level 1
  - Scopus rating (2014): SJR 1.719 SNIP 1.36 CiteScore 3.11
  - BFI (2013): BFI-level 1
  - Scopus rating (2013): SJR 1.512 SNIP 1.488 CiteScore 2.86
  - ISI indexed (2013): ISI indexed yes
  - Web of Science (2013): Indexed yes
  - BFI (2012): BFI-level 1
  - Scopus rating (2012): SJR 1.648 SNIP 1.192 CiteScore 2.81
  - ISI indexed (2012): ISI indexed yes
  - BFI (2011): BFI-level 1
  - Scopus rating (2011): SJR 1.363 SNIP 1.231 CiteScore 2.73
  - ISI indexed (2011): ISI indexed yes
  - Web of Science (2011): Indexed yes
  - BFI (2010): BFI-level 1
  - Scopus rating (2010): SJR 1.485 SNIP 1.137
  - Web of Science (2010): Indexed yes
  - BFI (2009): BFI-level 1
  - Scopus rating (2009): SJR 1.31 SNIP 1.105
  - BFI (2008): BFI-level 2
  - Scopus rating (2008): SJR 1.521 SNIP 0.997
  - Web of Science (2008): Indexed yes
  - Scopus rating (2007): SJR 1.165 SNIP 1.061
  - Web of Science (2007): Indexed yes
  - Scopus rating (2006): SJR 0.981 SNIP 0.927
Perception and neural representation of tones in conditions of masking release

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Egger, K. (Intern), Epp, B. (Intern)
Pages: 238
Publication date: 2013
Conference: 36th Annual Midwinter Meeting of the Association for Research in Otolaryngology, Baltimore, MD, United States, 16/02/2013 - 16/02/2013
Main Research Area: Technical/natural sciences

Publication information
Journal: ARO Abstracts
Volume: 36
Original language: English
Electronic versions:
aro_2013_Egger_Epp_Abstract.pdf
Source: dtu
Source-ID: u::8888
Publication: Research - peer-review › Journal article – Annual report year: 2013

Playful Interaction with Voice Sensing Modular Robots

This paper describes a voice sensor, suitable for modular robotic systems, which estimates the energy and fundamental frequency, $F_0$, of the user's voice. Through a number of example applications and tests with children, we observe how the voice sensor facilitates playful interaction between children and two different robot configurations. In future work, we will investigate if such a system can motivate children to improve voice control and explore how to extend the sensor to detect emotions in the user's voice.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Automation and Control, Centre for Playware, Technical University of Denmark
Authors: Heesche, B. (Ekstern), MacDonald, E. (Intern), Fogh, R. (Intern), Pacheco, M. (Intern), Christensen, D. J. (Intern)
Pages: 180–189
Publication date: 2013
Conference: ICSR 2013, Bristol, United Kingdom, 27/10/2013 - 27/10/2013
Main Research Area: Technical/natural sciences
Title of host publication: Proceedings of ICSR 2013
Publisher: Springer
ISBN (Print): 978-3-319-02674-9
Series: Lecture Notes in Artificial Intelligence
Volume: 8239
ISSN: 0302-9743
Predicting speech intelligibility in conditions with nonlinearly processed noisy speech

The speech-based envelope power spectrum model (sEPSM; [1]) was proposed in order to overcome the limitations of the classical speech transmission index (STI) and speech intelligibility index (SII). The sEPSM applies the signal-to-noise ratio in the envelope domain (SNRenv), which was demonstrated to successfully predict speech intelligibility in conditions with nonlinearly processed noisy speech, such as processing with spectral subtraction. Moreover, a multiresolution version (mr-sEPSM) was demonstrated to account for speech intelligibility in various conditions with stationary and fluctuating interferers [2]. However, the model fails in the case of phase jitter distortion, in which the spectral structure of speech is affected but the temporal envelope is maintained. This suggests that an across audio-frequency mechanism is required to account for this distortion. It is demonstrated that a measure of the across audio-frequency variance at the output of the modulation-frequency selective process in the model is sufficient to account for the phase jitter distortion. Thus, a joint spectro-temporal modulation analysis, as proposed in [3], does not seem to be required. The results are consistent with concepts from computational auditory scene analysis and further support the hypothesis that the SNRenv is a powerful metric for speech intelligibility prediction.

Reliability of procedures used for scaling loudness

In this study, 16 normally-hearing listeners judged the loudness of 1000-Hz sinusoids using magnitude estimation (ME), magnitude production (MP), and categorical loudness scaling (CLS). Listeners in each of four groups completed the loudness scaling tasks in a different sequence on the first visit (ME, MP, CLS; MP, ME, CLS; CLS, ME, MP; CLS, MP, ME), and the order was reversed on the second visit.

This design made it possible to compare the reliability of estimates of the slope of the loudness function across procedures in the same listeners. The ME data were well fitted by an inflected exponential (INEX) function, but a modified power law was used to obtain slope estimates for both ME and MP. ME and CLS were more reliable than MP. CLS results were consistent across groups, but ME and MP results differed across groups in a way that suggested influence of experience with CLS. Although CLS results were the most reproducible, they do not provide direct information about the slope of the loudness function because the numbers assigned to CLS categories are arbitrary. This problem can be corrected by using data from the other procedures to assign numbers that are proportional to loudness. [Supported by NIH]
Sequential dependencies in magnitude scaling of loudness

Ten normally hearing listeners used a programmable sone-potentiometer knob to adjust the level of a 1000-Hz sinusoid to match the loudness of numbers presented to them in a magnitude production task. Three different power-law exponents (0.15, 0.30, and 0.60) and a log-law with equal steps in dB were used to program the sone-potentiometer. The knob settings systematically influenced the form of the loudness function. Time series analysis was used to assess the sequential dependencies in the data, which increased with increasing exponent and were greatest for the log-law. It would be possible, therefore, to choose knob properties that minimized these dependencies. When the sequential dependencies were removed from the data, the slope of the loudness functions did not change, but the variability decreased. Sequential dependencies were only present when the level of the tone on the previous trial was higher than on the current trial. According to the attention band hypothesis [Green and Luce, 1974, Perception & Psychophysics] these dependencies arise from a process similar to selective attention, but observations of rapid adaptation of neurons in the inferior colliculus based on stimulus level statistics [Dean et al., 2005, Nature Neuroscience] would also account for the data.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Boys Town National Research Hospital
Authors: Joshi, S. N. (Intern), Jesteadt, W. (Ekstern)
Number of pages: 7
Publication date: 2013
Conference: 21st International Congress on Acoustics, Montreal, Canada, 02/06/2013 - 02/06/2013
Main Research Area: Technical/natural sciences

Publication information
Journal: Meetings on Acoustics. Proceedings
Volume: 19
ISSN (Print): 1939-800X
Ratings:
Scopus rating (2016): SNIP 0.124 CiteScore 0.15
Scopus rating (2015): SJR 0.218 SNIP 0.147
Scopus rating (2014): SJR 0.188 SNIP 0.105
Scopus rating (2013): SJR 0.138 SNIP 0.243
ISI indexed (2013): ISI indexed no
Scopus rating (2012): SJR 0.144 SNIP 0.21
ISI indexed (2012): ISI indexed no
Scopus rating (2011): SJR 0.141 SNIP 0.313
ISI indexed (2011): ISI indexed no
Scopus rating (2010): SJR 0.163 SNIP 0.09
Scopus rating (2009): SJR 0.129 SNIP 0.106
Original language: English
Source: dtu
Source-ID: u::7797
Publication: Research - peer-review › Conference article – Annual report year: 2013
Simulating psychophysical tuning curves in listeners with dead regions

Objective: This study investigates the relation between diagnosis of dead regions based on the off-frequency psychophysical tuning curve (PTC) tip and the frequency and level of the probe tone. Design: A previously developed functional model of auditory processing was used to simulate the complete loss of inner hair cells (IHC), dysfunction of outer hair cells (OHC), complete loss of IHCs in combination with OHC dysfunction, and IHC insensitivity. The model predictions were verified through comparison with experimental data. Study sample: This study compares PTC data of five normal-hearing listeners and six hearing-impaired listeners with model-simulated PTC data. Results: It was shown that OHC activity and IHC insensitivity may significantly alter the shift of PTC tips with increasing probe level. Conclusions: Model results suggest that OHC activity and IHC insensitivity can change the outcome of dead region diagnosis using PTCs. Supplementary to PTC dead region diagnostic information, model results may provide additional information regarding the edge frequency of a dead region and OHC function.
Sound-field reconstruction performance of a mixed-order Ambisonics microphone array

Recently, there has been increasing interest in using spherical microphone arrays for spatial audio recordings. Accurate recordings are important for a range of applications, from virtual sound environments for hearing research through to the evaluation of communication devices, such as hearing instruments and mobile phones. Previously, a mixed-order Ambisonics (MOA) approach was proposed to improve the horizontal spatial resolution of spherical arrays. This was achieved by increasing the number of microphones near the horizontal plane while keeping the total number of transducers fixed. The approach is motivated by the fact that human spatial hearing is most acute in the horizontal plane.

This study presents simulations of the performance of an MOA rigid-sphere microphone array, and its robustness to variations in microphone characteristics. Specifications of a commercially available microphone were used to simulate self-noise, sensitivity, and phase response variations between the microphones. To quantify the reconstruction error and the "sweet area" as a function of source elevation, the reconstructed sound field based on a simulated array measurement was compared to the reference sound field for both horizontal and elevated sources. It is expected that the MOA approach results in a larger sweet area for mid to high frequencies for horizontal sources.

General information

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Acoustic Technology
Authors: Marschall, M. (Intern), Chang, J. (Intern)
Number of pages: 9
Publication date: 2013
Conference: 21st International Congress on Acoustics, Montreal, Canada, 02/06/2013 - 02/06/2013
Main Research Area: Technical/natural sciences

Publication information

Journal: Meetings on Acoustics. Proceedings
Volume: 19
ISSN (Print): 1939-800X
Ratings:
Scopus rating (2016): SNIP 0.124 CiteScore 0.15
Scopus rating (2015): SJR 0.218 SNIP 0.147
Scopus rating (2014): SJR 0.188 SNIP 0.105
Scopus rating (2013): SJR 0.138 SNIP 0.243
ISI indexed (2013): ISI indexed no
Scopus rating (2012): SJR 0.144 SNIP 0.21
ISI indexed (2012): ISI indexed no
Scopus rating (2011): SJR 0.141 SNIP 0.313
ISI indexed (2011): ISI indexed no
Scopus rating (2010): SJR 0.163 SNIP 0.09
Scopus rating (2009): SJR 0.129 SNIP 0.106
Original language: English
DOIs:
10.1121/1.4800859
Sound Source Distance Estimation in Rooms based on Statistical Properties of Binaural Signals

A novel method for the estimation of the distance of a sound source from binaural speech signals is proposed. The method relies on several statistical features extracted from such signals and their binaural cues. Firstly, the standard deviation of the difference of the magnitude spectra of the left and right binaural signals is used as a feature for this method. In addition, an extended set of additional statistical features that can improve distance detection is extracted from an auditory front-end which models the peripheral processing of the human auditory system. The method incorporates the above features into two classification frameworks based on Gaussian mixture models and Support Vector Machines and the relative merits of those frameworks are evaluated. The proposed method achieves distance detection when tested in various acoustical environments and performs well in unknown environments. Its performance is also compared to an existing binaural distance detection method.
Spectral information for detection of acoustic time to arrival

The exponential increase of intensity for an approaching sound source provides salient information for a listener to make judgments of time to arrival (TTA). Specifically, a listener will experience a greater rate of increasing intensity for higher than for lower frequencies during a sound source’s approach. To examine the relative importance of this spectral information, listeners were asked to make judgments about the arrival times of nine 1-octave-band sound sources (the bands were consecutive, nonoverlapping single octaves, ranging from 40–80 Hz to ~10–20 kHz). As is typical in TTA tasks, listeners tended to underestimate the arrival time of the approaching sound source. In naturally occurring and independently manipulated amplification curves, bands with center frequencies between 120 and 250 Hz caused the least underestimation, and bands with center frequencies between 2000 and 7500 Hz caused the most underestimation. This spectral influence appears to be related to the greater perceived urgency of higher-frequency sounds.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, William Paterson University
Authors: Gordon, M. S. (Ekstern), Russo, F. A. (Ekstern), MacDonald, E. (Intern)
Pages: 738–750
Publication date: 2013
Main Research Area: Technical/natural sciences

Publication information
Journal: Attention, Perception & Psychophysics
Volume: 75
ISSN (Print): 1943-3921
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.77 SJR 1.174 SNIP 0.904
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 1.378 SNIP 0.925 CiteScore 1.88
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 1.399 SNIP 1.093 CiteScore 2.27
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 1.667 SNIP 1.18 CiteScore 2.14
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 1.495 SNIP 1.078 CiteScore 2.03
ISI indexed (2012): ISI indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 1.585 SNIP 1.169 CiteScore 1.86
ISI indexed (2011): ISI indexed no
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 1.138 SNIP 0.939
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 1.322 SNIP 1.009
Spectral integration of interaural time differences in auditory localization

This study investigates how the auditory system integrates spatial information across frequency. In experiment 1, discrimination thresholds for interaural time differences (ITDs) were measured as a function of both reference ITD and center frequency (CF) of noises with bandwidth of one ERB. In addition, discrimination thresholds were also measured as a function of CF for different values of interaural coherence (IC) typical of sounds in realistic acoustic environments. For both high ICs and small reference ITDs, discrimination thresholds were lowest for CFs between 700 and 1000 Hz. For smaller ICs and larger reference ITDs, this dominance region shifted towards lower CFs. A conceptual localization model was developed that used the variance of the ITD thresholds to optimally weight the contribution of the individual frequency bands before spectral integration. In experiment 2, the model was tested by asking listeners to align a broadband noise signal with an ITD that was fixed across frequency onto a broadband noise target with different ITDs in individual 1 ERB-wide subbands. The results were consistent with both the model predictions and the shift of dominance range observed in experiment one.
Spectral Integration of Interaural Time Differences in Auditory Localization

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Macquarie University
Authors: Le Goff, N. (Intern), Buchholz, J. M. (Ekstern), Dau, T. (Intern)
Number of pages: 1
Publication date: 2013
Event: Poster session presented at 21st International Congress on Acoustics, Montreal, Canada.
Main Research Area: Technical/natural sciences
Electronic versions:
PosterASAICA2013_NLG

Spectrogram inversion and potential applications for hearing research
A common way of analyzing signals in a joint time-frequency domain is found in the spectrogram, which can be interpreted as a multi-channel envelope representation of the signal. The envelope cannot fully represent a signal because it only reflects slow changes in the amplitude of a signal and lacks information regarding its fast variations, the temporal fine structure (TFS). However, the main hypothesis explored in this thesis is that a spectrogram could be a faithful representation of a signal, that is, TFS information could be recovered by across-channel comparison of envelopes. Based on this consideration, an approach for spectrogram inversion was proposed: time-domain signals were recovered from spectrograms computed using both inner hair-cell envelope (i.e., traditional half-wave rectification followed by low-pass filtering) and Hilbert envelope definitions. The high accuracy of the inversion scheme (as measured by root mean square error and spectral convergence) implies that the main hypothesis holds true for the designs chosen. Two practical applications of this result were then presented. (1) Spectrograms that are computed using the inner hair-cell (IHC) envelope definition are a reasonable model of the signal processing performed by the human cochlea. The robustness of the reconstruction from such spectrograms with regards to the properties of the cochlear model showed that, for previously documented IHC models as well as for more restrictive conditions, the TFS-related information is retained by the (modeled) cochlear processing even at high audio frequencies. (2) Using the inversion framework, it is possible to manipulate signals in the modulation domain, while preserving their long-term power spectra. Thus, this enabled the creation of mixtures of speech and noise where the signal-to-noise ratio in the envelope domain (SNRenv) was directly controlled. Behavioral measures of the intelligibility for such mixtures were compared to predictions from a model of speech intelligibility. Conditions where noise was processed led to modest intelligibility improvements for increased SNRenv, providing v

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Decorsière, R. J. B. (Intern), Dau, T. (Intern), Søndergaard, P. L. (Intern), MacDonald, E. (Intern)
Publication date: 2013

Publication information
Publisher: Technical University of Denmark, Department of Electrical Engineering
Original language: English
Main Research Area: Technical/natural sciences
Electronic versions:
PhD_RD_final.pdf
Source: PublicationPreSubmission
Source-ID: 105128951
Publication: Research › Ph.D. thesis – Annual report year: 2013
Speech production in amplitude-modulated noise

The Lombard effect refers to the phenomenon where talkers automatically increase their level of speech in a noisy environment. While many studies have characterized how the Lombard effect influences different measures of speech production (e.g., F0, spectral tilt, etc.), few have investigated the consequences of temporally fluctuating noise. In the present study, 20 talkers produced speech in a variety of noise conditions, including both steady-state and amplitude-modulated white noise. While listening to noise over headphones, talkers produced randomly generated five word sentences. Similar to previous studies, talkers raised the level of their voice in steady-state noise. While talkers also increased the level of their voice in amplitude-modulated noise, the increase was not as large as that observed in steady-state noise. Importantly, for the 2 and 4 Hz amplitude-modulated noise conditions, talkers altered the timing of their utterances, reducing the energetic overlap with the masker by approximately 2%. However, for the 1 Hz amplitude-modulated condition, talkers increased the overlap by approximately 4%. Overall, the results demonstrate that talkers are sensitive to the temporal aspects of noisy environments and will alter their speech accordingly.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Jade Hochschule, Oldenburg
Authors: Macdonald, E. N. (Intern), Raufer, S. (Ekstern)
Number of pages: 7
Pages: 060149-
Publication date: 2013
Conference: 21st International Congress on Acoustics, Montreal, Canada, 02/06/2013 - 02/06/2013
Main Research Area: Technical/natural sciences

Publication information
Journal: Meetings on Acoustics. Proceedings
Volume: 19
ISSN (Print): 1939-800X
Ratings:
Scopus rating (2016): SNIP 0.124 CiteScore 0.15
Scopus rating (2015): SJR 0.218 SNIP 0.147
Scopus rating (2014): SJR 0.188 SNIP 0.105
Scopus rating (2013): SJR 0.138 SNIP 0.243
ISI indexed (2013): ISI indexed no
Scopus rating (2012): SJR 0.144 SNIP 0.21
ISI indexed (2012): ISI indexed no
Scopus rating (2011): SJR 0.141 SNIP 0.313
ISI indexed (2011): ISI indexed no
Scopus rating (2010): SJR 0.163 SNIP 0.09
Scopus rating (2009): SJR 0.129 SNIP 0.106
Original language: English
DOIs:
10.1121/1.4806308

Bibliographical note
Session 4aSCa: Auditory Feedback in Speech Production 1
Source: dtu
Source-ID: n:oai:DTIC-ART:pubmed/386404861::28587
Publication: Research - peer-review » Conference article – Annual report year: 2013

Temporal integration near threshold fine structure - The role of cochlear processing
The hearing thresholds of normal hearing listeners often show quasi-periodic variations when measured with a high frequency resolution. This hearing threshold fine structure is related to other frequency specific variations in the perception of sound such as loudness and amplitude modulated tones at low intensities. The detection threshold of a pulsed tone also depends not only on the pulse duration, but also on the position of its frequency within threshold fine structure. The present study investigates if psychoacoustical data on detection of a pulsed tone can be explained with a nonlinear and active transmission line cochlea model. The model was extended by including a temporal integrator which introduces a low-pass behavior of the data with different slopes of the predicted threshold curves, producing good agreement with the data. On the basis of the model simulations, it will be discussed to which extent temporal and spectral aspects contribute to the data.
The effect of cochlear nonlinearities on binaural masking level differences

Background
The binaural masking level difference (BMLD) has been shown to be constant (10−15dB) for masker spectrum levels from 70dB/Hz down to 30−40dB/Hz and to gradually decrease with lower levels (McFadden, 1968; Hall and Harvey, 1984). The decrease at low levels was larger in an asymmetric condition where the masker was attenuated in only one ear. McFadden predicted the data by assuming that an external and an internal noise would interaurally decorrelate the internal representations of the stimuli. In the present study, the role of nonlinear cochlear processing and asymmetric masker level on the BMLD was investigated using an equalization-cancelation (EC) based binaural model framework.

Methods
The BMLD was measured for 500−Hz target tones presented in 3−kHz−wide maskers. BMLDs were obtained as a function of masker level in one symmetric and two asymmetric masker conditions: (i) NoSπ: thresholds were measured for masker spectrum levels between 50 and -10dB/Hz, with the same level at both ears; (ii) No′Sπ′50: same as first condition but the masker was attenuated in one ear only and fixed at 50dB/Hz in the non−attenuated ear; (iii) No′SΠ′20: same as second condition but with a masker level of 20dB/Hz in the non−attenuated ear. An EC based binaural model with a frontend including nonlinear peripheral processing (Jepsen et al., 2011) was used to predict these results.

Results
The BMLD obtained in the No′Sπ′50 condition was smaller than that obtained in the NoSπ condition at all masker levels between 50 and -10dB/Hz. The difference in BMLD between the two conditions gradually increased with decreasing masker level from 50 to 20dB/Hz and remained constant between 20 and -10dB/Hz. The proposed model could account for these data. A model analysis suggested that the increase in BMLD difference between the No′Sπ′50 and NoSπ conditions results from a decrease in interaural correlation at the output of the periphery, and is a consequence of the cochlear nonlinearity at levels between 20 and 50dB/Hz. For levels below 20dB/Hz, cochlear processing becomes linear and the model predicts that the difference in BMLD between the symmetric and the two asymmetric conditions is constant, in line with the experimental data.

Conclusion
A model was proposed to account for the effect of level asymmetry on BMLD. The modeling results suggest that cochlear nonlinearities affect the analysis of binaural cues at higher processing stages such that signals carrying interaural level differences can become interaurally decorrelated.
The effect of compression on tuning estimates in a simple nonlinear auditory filter model

Behavioral experiments using auditory masking have been used to characterize frequency selectivity, one of the basic properties of the auditory system. However, due to the nonlinear response of the basilar membrane, the interpretation of these experiments may not be straightforward. Specifically, there is evidence that human frequency-selectivity estimates depend on whether an iso-input or an iso-response measurement paradigm is used (Eustaquio-Martin et al., 2011). This study presents simulated tuning estimates using a simple compressive auditory filter model, the bandpass nonlinearity (BPNL), which consists of a compressor between two bandpass filters. The BPNL forms the basis of the dual-resonance nonlinear (DRNL) filter that has been used in a number of modeling studies. The location of the nonlinear element and its effect on estimated tuning in the two measurement paradigms was investigated. The results show that compression leads to (i) a narrower tuning estimate in the iso-response paradigm when a compressor precedes a filter, and (ii) a wider tuning estimate in the iso-input paradigm when a compressor follows a filter. The results imply that if the DRNL presents a valid cochlear model, then compression alone may explain a large part of the behaviorally observed differences in tuning between simultaneous and forward-masking conditions.

General information

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Marschall, M. (Intern), MacDonald, E. (Intern), Dau, T. (Intern)
Number of pages: 5
Publication date: 2013
Conference: 21st International Congress on Acoustics, Montreal, Canada, 02/06/2013 - 02/06/2013
Main Research Area: Technical/natural sciences

Publication information

Journal: Meetings on Acoustics. Proceedings
Volume: 19
ISSN (Print): 1939-800X
Ratings:
Scopus rating (2016): SNIP 0.124 CiteScore 0.15
Scopus rating (2015): SJR 0.218 SNIP 0.147
Scopus rating (2014): SJR 0.188 SNIP 0.105
Scopus rating (2013): SJR 0.138 SNIP 0.243
ISI indexed (2013): ISI indexed no
Scopus rating (2012): SJR 0.144 SNIP 0.21
ISI indexed (2012): ISI indexed no
Scopus rating (2011): SJR 0.141 SNIP 0.313
ISI indexed (2011): ISI indexed no
Scopus rating (2010): SJR 0.163 SNIP 0.09
Scopus rating (2009): SJR 0.129 SNIP 0.106
Original language: English
Electronic versions:
Marschall et al. (2013) Effect of compression on tuning estimates in a simple nonlinear auditory filter model.pdf
DOIs:
10.1121/1.4799637

Bibliographical note

The attached file is the published conference proceeding for the International Congress on Acoustics: Proceedings of Meetings on Acoustics, Vol. 19, 050109 (2013)
Source: dtu
Source-ID: u::7555
Publication: Research - peer-review › Conference article – Annual report year: 2013

The effect of interaural-level-difference fluctuations on the externalization of sound

Real-world sound sources are usually perceived as externalized and thus properly localized in both direction and distance. This is largely due to (1) the acoustic filtering by the head, torso, and pinna, resulting in modifications of the signal spectrum and thereby a frequency-dependent shaping of interaural cues and (2) interaural cues provided by the reverberation inside an enclosed space. This study first investigated the effect of room reverberation on the spectro-temporal behavior of interaural level differences (ILDs) by analyzing dummy-head recordings of speech played at different distances in a standard listening room. Next, the effect of ILD fluctuations on the degree of externalization was investigated in a psychoacoustic experiment performed in the same listening room. Individual binaural impulse responses were used to simulate a distant sound source delivered via headphones. The ILDs were altered using a gammatone filter.
filterbank for analysis and resynthesis, where the envelopes of the left and right-ear signals were modified such that the naturally occurring fluctuations of the ILDs were restricted. This manipulation reduced the perceived degree of externalization. This was consistent with the analysis of short-term ILDs at different distances showing that a decreased distance to the sound source also reduced the ILD fluctuations. © 2013 Acoustical Society of America.
The influence of masker type on early reflection processing and speech intelligibility (L)
Arweiler and Buchholz [J. Acoust. Soc. Am. 130, 996-1005 (2011)] showed that, while the energy of early reflections (ERs) in a room improves speech intelligibility, the benefit is smaller than that provided by the energy of the direct sound (DS). In terms of integration of ERs and DS, binaural listening did not provide a benefit from ERs apart from a binaural energy summation, such that monaural auditory processing could account for the data. However, a diffuse speech shaped noise (SSN) was used in the speech intelligibility experiments, which does not provide distinct binaural cues to the auditory system. In the present study, the monaural and binaural benefit from ERs for speech intelligibility was investigated using three directional maskers presented from 90° azimuth: a SSN, a multi-talker babble, and a reversed two-talker masker. For normal-hearing as well as hearing-impaired listeners, the directional and/or fluctuating (speech) maskers produced a similar benefit from ERs as obtained with the diffuse SSN, suggesting a monaural integration of the ERs and the DS for both types of maskers. © 2013 Acoustical Society of America

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Centre for Applied Hearing Research
Authors: Arweiler, I. (Intern), Buchholz, J. M. (Intern), Dau, T. (Intern)
Pages: 13-16
Publication date: 2013
Main Research Area: Technical/natural sciences

Publication information
Volume: 133
Issue number: 1
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
The relation between perceived apparent source width and interaural cross-correlation in sound reproduction spaces with low reverberation

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Käsbach, J. (Intern), Marschall, M. (Intern), Epp, B. (Intern), Dau, T. (Intern)
Number of pages: 4
Publication date: 2013
The Relationship Between Background Noise Envelope Power and Speech Intelligibility in Adverse Conditions

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Jørgensen, S. (Intern), Decorsiere, R. J. B. (Intern), MacDonald, E. (Intern), Dau, T. (Intern)
Number of pages: 1
Publication date: 2013
Event: Poster session presented at 36th Annual Midwinter Meeting of the Association for Research in Otolaryngology, Baltimore, MD, United States.
Main Research Area: Technical/natural sciences
Electronic versions:
The_relationship.pdf

Bibliographical note
The attached PDF that includes the abstract and poster.
Source: dtu
Source-ID: u::7025
Publication: Research - peer-review › Poster – Annual report year: 2013

The role of across-frequency envelope processing for speech intelligibility
Speech intelligibility models consist of a preprocessing part that transforms the stimuli into some internal (auditory) representation, and a decision metric that quantifies effects of transmission channel, speech interferers, and auditory processing on the speech intelligibility. Here, two recent speech intelligibility models, the spectro-temporal modulation index [STMI; Elhilali et al. (2003)] and the speech-based envelope power spectrum model [sEPSM; Jørgensen and Dau (2011)] were evaluated in conditions of noisy speech subjected to reverberation, and to nonlinear distortions through either a phase jitter process or noise reduction via spectral subtraction. The contributions of the individual preprocessing stages in the models and the role of the decision metrics were analyzed in the different experimental conditions. It is demonstrated that an explicit across-frequency envelope processing stage, as assumed in the STMI, together with the metric based on the envelope power signal-to-noise ratio, as assumed in the sEPSM, are required to account for all three conditions. However, a simple weighting of the across-frequency variance of the modulation power at the output of the (purely temporal) modulation filterbank is assumed to be sufficient to describe the data, i.e., a joint two-dimensional modulation filterbank might not be required.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Chabot-Leclerc, A. (Intern), Jørgensen, S. (Intern), Dau, T. (Intern)
Pages: 3391-3391
Publication date: 2013
Conference: 21st International Congress on Acoustics, Montreal, Canada, 02/06/2013 - 02/06/2013
Main Research Area: Technical/natural sciences

Publication information
Volume: 133
Issue number: 5
Article number: 2pSCb20
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
The role of across-frequency envelope processing for speech intelligibility

Speech intelligibility models consist of a preprocessing part that transforms the stimuli into some internal (auditory) representation, and a decision metric that quantifies effects of transmission channel, speech interferers, and auditory processing on the speech intelligibility. Here, two recent speech intelligibility models, the spectro-temporal modulation index (STMI; Elhilali et al., 2003) and the speech-based envelope power spectrum model (sEPSM; Jørgensen and Dau, 2011) were evaluated in conditions of noisy speech subjected to reverberation, and to nonlinear distortions through either a phase jitter process or noise reduction via spectral subtraction. The contributions of the individual preprocessing stages in the models and the role of the decision metrics were analyzed in the different experimental conditions. It is demonstrated that an explicit across-frequency envelope processing stage, as assumed in the STMI, together with the metric based on the envelope power signal-to-noise ratio, as assumed in the sEPSM, are required to account for all three conditions. However, a simple across audio-frequency mechanism combined with a purely temporal modulation filterbank is assumed to be sufficient to describe the data, i.e., a joint two-dimensional modulation filterbank might not be required.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Chabot-Leclerc, A. (Intern), Jørgensen, S. (Intern), Dau, T. (Intern)
Number of pages: 8
Publication date: 2013
Conference: 21st International Congress on Acoustics, Montreal, Canada, 02/06/2013 - 02/06/2013
Main Research Area: Technical/natural sciences

Publication information
Journal: Meetings on Acoustics. Proceedings
Volume: 19
ISSN (Print): 1939-800X
Ratings:
Scopus rating (2016): SNIP 0.124 CiteScore 0.15
Scopus rating (2015): SJR 0.218 SNIP 0.147
Scopus rating (2014): SJR 0.188 SNIP 0.105
Scopus rating (2013): SJR 0.138 SNIP 0.243
ISI indexed (2013): ISI indexed no
Scopus rating (2012): SJR 0.144 SNIP 0.21
ISI indexed (2012): ISI indexed no
Scopus rating (2011): SJR 0.141 SNIP 0.313
ISI indexed (2011): ISI indexed no
Scopus rating (2010): SJR 0.163 SNIP 0.09
Scopus rating (2009): SJR 0.129 SNIP 0.106
Original language: English
DOIs:
10.1121/1.4799026
Publication: Research - peer-review › Conference article – Annual report year: 2013

The role of high-frequency envelope fluctuations for speech masking release

The speech-based envelope power spectrum model (sEPSM; Jørgensen and Dau, 2011; Jørgensen et al., 2013) was shown to successfully predict speech intelligibility in conditions with stationary and fluctuating interferers, reverberation, and spectral subtraction. The key element in the model was the multi-resolution estimation of the signal-to-noise ratio in the envelope domain (SNRenv) at the output of a modulation filterbank. The simulations suggested that mainly modulation filters centered in the range from 1-8 Hz contribute to speech intelligibility in the case of stationary maskers whereas modulation filters tuned to frequencies above 16 Hz might be important in the case of fluctuating maskers. In the present study, the role of high-frequency envelope fluctuations for speech masking release was further investigated in conditions of speech-on-speech masking. Simulations were compared to various measured data from normal-hearing listeners (Festen and Plomp, 1990; Christiansen et al., 2013). The results support the hypothesis that high-frequency envelope fluctuations (>30 Hz) are essential for speech intelligibility in conditions with speech interferers. While the sEPSM reflects effects of energetic and modulation masking in speech intelligibility, the remaining unexplored effect in some conditions may be attributed to, and defined as, "informational masking".

General information
State: Published
The role of high-frequency envelope fluctuations for speech masking release.
The speech-based envelope power spectrum model [sEPSM; Jørgensen and Dau (2011), Jørgensen et al. (2013)] was shown to successfully predict speech intelligibility in conditions with stationary and fluctuating interferers, reverberation, and spectral subtraction. The key element in the model was the multi-resolution estimation of the signal-to-noise ratio in the envelope domain (SNRenv) at the output of a modulation filterbank. The simulations suggested that mainly modulation filters centered in the range from 1 to 8 Hz contribute to speech intelligibility in the case of stationary maskers whereas modulation filters tuned to frequencies above 16 Hz might be important in the case of fluctuating maskers. In the present study, the role of high-frequency envelope fluctuations for speech masking release was further investigated in conditions of speech-on-speech masking. Simulations were compared to various measured data from normal-hearing and hearing-impaired listeners [Festen and Plomp (1990), Christiansen et al. (2013)]. The results support the hypothesis that high-frequency envelope fluctuations (>30 Hz) are essential for speech intelligibility in conditions with speech interferers. While the sEPSM reflects effects of energetic and modulation masking in speech intelligibility, the remaining unexplored effect in some conditions may be attributed to, and defined as, “information masking.”

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Jørgensen, S. (Intern), Dau, T. (Intern)
Number of pages: 1
Pages: 3391
Publication date: 2013
Conference: 21st International Congress on Acoustics, Montreal, Canada, 02/06/2013 - 02/06/2013
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of the Acoustical Society of America
Volume: 133
Issue number: 5
ISSN (Print): 0001-4966
Ratings: BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
Validation of a Virtual Sound Environment System for Hearing Aid Testing

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Oticon A/S
Authors: Cubick, J. (Intern), Favrot, S. E. (Intern), Minnaar, P. (Ekstern), Dau, T. (Intern)
Pages: 2388-2390
Publication date: 2013

Host publication information
Title of host publication: Proceedings of the International Conference on Acoustics
Main Research Area: Technical/natural sciences
Conference: AIA-DAGA 2013 Conference on Acoustics, Merano, Italy, 18/05/2013 - 18/05/2013
Source: dtu
Source-ID: u::7894
Publication: Research - peer-review › Article in proceedings – Annual report year: 2013

A computational model of auditory stream segregation based on a temporal coherence analysis

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Christiansen, S. K. (Intern), Jepsen, M. L. (Intern), Dau, T. (Intern)
Number of pages: 2
Publication date: 2012

Host publication information
Title of host publication: Proceedings of the Deutsche Gesellschaft für Akustik
Main Research Area: Technical/natural sciences
Conference: 38th German Annual Conference On Acoustics, Darmstadt, Germany, 19/03/2012 - 19/03/2012
Publication: Research - peer-review › Article in proceedings – Annual report year: 2012

A physiologically inspired model of auditory stream segregation based on a temporal coherence analysis

The ability to perceptually separate acoustic sources and focus one’s attention on a single source at a time is essential for our ability to use acoustic information. In this study, a physiologically inspired model of human auditory processing [M. L. Jepsen and T. Dau, J. Acoust. Soc. Am. 124, 422-438, (2008)] was used as a front end of a model for auditory stream segregation. A temporal coherence analysis [M. Elhilali, C. Ling, C. Micheyl, A. J. Oxenham and S. Shamma, Neuron. 61, 317-329, (2009)] was applied at the output of the preprocessing, using the coherence across tonotopic channels to group activity across frequency. Using this approach, the described model is able to quantitatively account for classical streaming phenomena relying on frequency separation and tone presentation rate, such as the temporal coherence boundary and the fission boundary [L. P. A. S. van Noorden, doctoral dissertation, Institute for Perception Research, Eindhoven, NL, (1975)]. The same model also accounts for the perceptual grouping of distant spectral components in the case of synchronous presentation. The most essential components of the front-end and back-end processing in the framework of the presented model are analysed and future perspectives discussed.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Christiansen, S. K. (Intern), Jepsen, M. L. (Intern), Dau, T. (Intern)
Number of pages: 7
Publication date: 2012

Host publication information
Title of host publication: Proceedings of Meetings on Acoustics
Volume: 15
Publisher: Acoustical Society of America
Main Research Area: Technical/natural sciences
A Preliminary Study of Individual Responses to Real-Time Pitch and Formant Perturbations

Previous studies have demonstrated a wide range in individuals’ compensations in response to real-time alterations of the auditory feedback of both pitch and formant frequencies. One potential source of this variability may be individual differences in the relative weighting of auditory and somatosensory feedback. The present study examined this variability by comparing individuals’ compensations during two perturbation conditions: a pitch shift (+200 cents) and a formant shift (F1 +200 Hz, F2 -250 Hz). While no significant correlation was found between the two perturbation conditions, a modest correlation between compensations in pitch and formant frequency was observed within the pitch perturbation condition.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Queen's University
Authors: MacDonald, E. (Intern), Munhall, K. G. (Ekstern)
Pages: 32-35
Publication date: 2012

Behavioral and objective measures of the precedence effect

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Bianchi, F. (Intern), Verhulst, S. (Intern), Dau, T. (Intern)
Publication date: 2012
Event: Poster session presented at 35th MidWinter Meeting of the Association for Research in Otolaryngology, San Diego, CA, United States.
Main Research Area: Technical/natural sciences
Electronic versions:
FedericaBianchi_PosterARO.pdf

Relations
Activities:
35th MidWinter Meeting of the Association for Research in Otolaryngology
Publication: Research - peer-review › Poster – Annual report year: 2012

Blind estimation of the number of speech source in reverberant multisource scenarios based on binaural signals

In this paper, we present a new approach for estimating the number of active speech sources in the presence of interfering noise sources and reverberation. First, a binaural front-end is used to detect the spatial positions of all active sound sources, resulting in a binary mask for each candidate position. Then, each candidate position is characterized by a set of features. In addition to exploiting the overall spectral shape, a new set of mask-based features is proposed which aims at characterizing the pattern of the estimated binary mask. The decision stage for detecting a speech source is based on a support vector machine (SVM) classifier. A systematic analysis shows that the proposed algorithm is able to blindly determine the number and the corresponding spatial positions of speech sources in multisource scenarios and generalizes well to unknown acoustic conditions.

General information
Comodulation Is a Stronger Binding Cue Than the Common Fate of Frequency Swept Components

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Verhey, J. L. (Ekstern), Winter, I. M. (Ekstern), Epp, B. (Intern)
Publication date: 2012
Event: Abstract from 35th MidWinter Meeting of the Association for Research in Otolaryngology, San Diego, CA, United States.
Main Research Area: Technical/natural sciences
Publication: Research › Conference abstract for conference – Annual report year: 2012

Control of sound fields with a circular double-layer array of loudspeakers
This investigation is concerned with generating a controlled sound field for listeners inside a circular array of loudspeakers without disturbing people outside the array. Ideally this configuration would have the advantage that reflections from the surroundings would be of no concern. Inspired by the Kirchhoff-Helmholtz integral theorem a double-layer array of loudspeakers is used. Several solution methods are suggested and examined with computer simulations: pure contrast control, pure pressure matching, and a weighted combination. In order to compare the performance of the methods two performance indices are used, i) the ratio of the sound energy in the listening zone to the sound energy in the quiet zone, and ii) a normalised measure of the deviations between the desired and the generated sound field in the listening zone. The best compromise is obtained with the method that combines pure contrast control with a pressure matching technique.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Acoustic Technology
Authors: Chang, J. (Intern), Jacobsen, F. (Intern)
Number of pages: 12
Publication date: 2012

Host publication information
Title of host publication: Proceedings of Inter-Noise 2012
Publisher: Institute of Noise Control Engineering
BFI conference series: Inter-Noise (5010071)
Main Research Area: Technical/natural sciences
Conference: 41st International Congress and Exposition on Noise Control Engineering, New York City, NY, United States, 19/08/2012 - 19/08/2012
Electronic versions:
Control_of_sound_fields.pdf
Source: dtu
Source-ID: u::7053
Publication: Research - peer-review › Article in proceedings – Annual report year: 2012

Effects of diotic fringes on interaural disparity detection (L)
Detection thresholds were measured for interaural time differences (ITDs) and interaural level differences (ILDs) that were carried by probe segments embedded in otherwise diotic broadband noise (fringe). The duration of the probe was varied between 5 and 200 ms, and the duration of the fringe was between 5 and 100 ms. Consistent with results of Akeroyd and Bernstein [(2001). J. Acoust. Soc. Am. 110, 2516-2526], it was found that a 5-ms fringe placed before a 5-ms probe
(forward fringe) led to a larger threshold elevation than a 5-ms fringe placed after the probe (backward fringe). As suggested by Akeroyd and Bernstein, this effect was accounted for by a model providing an onset emphasis of a factor of 2. In contrast, for longer probe and fringe durations, which have not been tested before, a backward fringe had a stronger effect than a forward fringe. This surprising effect was accounted for by an extended model that provided an offset emphasis of a factor of 11 for a 50-ms probe and a 100-ms fringe.
Effects of Interaural Level and Time Differences on the Externalization of Sound

Distant sound sources in our environment are perceived as externalized and are thus properly localized in both direction and distance. This is due to the acoustic filtering by the head, torso, and external ears, which provides frequency dependent shaping of binaural cues, such as interaural level differences (ILDs) and interaural time differences (ITDs). Further, the binaural cues provided by reverberation in an enclosed space may also contribute to externalization. While these spatial cues are available in their natural form when listening to real-world sound sources, hearing-aid signal processing - such as wide dynamic range compression - affects the ILDs and thereby potentially reduces the perceived degree of externalization. In the present study, the effect of room reverberation on the spectro-temporal behavior of ILDs was investigated. This was done by analyzing speech played at different distances and recorded on a head-and-torso simulator in a standard IEC 268-13 listening room. Next, the effect of ILD fluctuations on the degree of externalization was investigated in a listening experiment with normal-hearing listeners. The experiment was performed in the same standard listening room and a distant speech source was simulated via headphones using individual binaural impulse responses. The speech signal was then processed such that the naturally occurring ILD fluctuations were compressed. This manipulation reduced the perceived degree of externalization in the listening experiment, which is consistent with the physical analysis that showed that a decreased distance to the sound source also reduced the fluctuations in ILDs.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, National Acoustic Laboratories
Authors: Dau, T. (Intern), Catic, J. (Intern), Santurette, S. (Intern), Buchholz, J. (Ekstern)
Publication date: 2012
Event: Abstract from 35th MidWinter Meeting of the Association for Research in Otolaryngology, San Diego, CA, United States.
Effects of interaural level differences on the externalization of sound

Distant sound sources in our environment are perceived as externalized and are thus properly localized in both direction and distance. This is due to the acoustic filtering by the head, torso, and external ears, which provides frequency-dependent shaping of binaural cues such as interaural level differences (ILDs) and interaural time differences (ITDs). In rooms, the sound reaching the two ears is further modified by reverberant energy, which leads to increased fluctuations in short-term ILDs and ITDs. In the present study, the effect of ILD fluctuations on the externalization of sound was investigated. A psychoacoustic experiment was performed in a standard IEC 268-13 listening room by normal-hearing listeners. Individual binaural room impulse responses were used to simulate a distant speech source delivered via headphones. The speech signal was then processed such that the naturally occurring fluctuations in the ILDs were compressed, while the ITDs were preserved. This manipulation reduced the perceived degree of externalization mainly for broadband and highpass filtered speech. In the case of lowpass filtered speech, the compression of ILD fluctuations did not affect externalization. Overall, for sounds that contain frequencies above about 1 kHz the ILD fluctuations were found to be an essential cue for externalization.
Efficient Algorithms for the Discrete Gabor Transform with a Long FIR Window

The Discrete Gabor Transform (DGT) is the most commonly used signal transform for signal analysis and synthesis using a linear frequency scale. The development of the Linear Time-Frequency Analysis Toolbox (LTFAT) has been based on a detailed study of many variants of the relevant algorithms. As a side result of these systematic developments of the subject, two new methods are presented here. Comparisons are made with respect to the computational complexity, and the running time of optimised implementations in the C programming language. The new algorithms have the lowest known computational complexity and running time when a long FIR window is used. The implementations are freely available for download. By summarizing general background information on the state of the art, this article can also be seen as a research survey, sharing with the readers experience in the numerical work in Gabor analysis.

General information

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Søndergaard, P. L. (Intern)
Pages: 456-470
Publication date: 2012
Main Research Area: Technical/natural sciences

Publication information

Journal: Journal of Fourier Analysis and Applications
Volume: 18
Issue number: 3
ISSN (Print): 1069-5869
Ratings:
BFI (2018): BFI-level 2
High-frequency complex pitch: a search for temporal cues and for a role of spectral indices

Harmonics in a complex tone are typically considered unresolved when they interact with neighboring harmonics in the cochlea and cannot be heard out separately. Recent studies have suggested that the low pitch evoked by unresolved high-frequency harmonics may be coded via temporal fine-structure cues. However, these conclusions rely on the assumptions that combination tones were properly masked and that the ability of listeners to hear out individual partials provides an adequate measure of resolvability. Those assumptions were tested by measuring the audibility of combination tones and their effects on pitch matches, the effects of relative component phases and of dichotic presentation, and
listeners' ability to hear out individual partials. The results confirmed that combination tones affected pitch, but pitch remained salient when they were masked. The lack of dependence of pitch on relative component phases or dichotic presentation provided no evidence in favor of temporal cues. Moreover, similar trends were observed between pitch salience and the listeners' ability to hear out individual partials. The results are consistent both with the use of place information and with a temporal code based on the combination of information across auditory channels.

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Santurette, S. (Intern), Dau, T. (Intern)
Pages: 3516
Publication date: 2012
Conference: Acoustics 2012 Hong Kong, Hong Kong, Hong Kong, 13/05/2012 - 13/05/2012
Main Research Area: Technical/natural sciences

**Publication information**

Volume: 131
Issue number: 4
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 0.848 SNIP 1.633
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 0.865 SNIP 1.647
Web of Science (2007): Indexed yes
Increased intensity discrimination thresholds in tinnitus subjects with a normal audiogram

Recent auditory brain stem response measurements in tinnitus subjects with normal audiograms indicate the presence of hidden hearing loss that manifests as reduced neural output from the cochlea at high sound intensities, and results from mice suggest a link to deafferentation of auditory nerve fibers. As deafferentation would lead to deficits in hearing performance, the present study investigates whether tinnitus patients with normal hearing thresholds show impairment in intensity discrimination compared to an audiometrically matched control group. Intensity discrimination thresholds were significantly increased in the tinnitus frequency range, consistent with the hypothesis that auditory nerve fiber deafferentation is associated with tinnitus.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Otto-von-Guericke University Magdeburg, University College London
Authors: Epp, B. (Intern), Hots, J. (Ekstern), Verhey, J. L. (Ekstern), Schaette, R. (Ekstern)
Pages: EL196-EL201
Publication date: 2012
Main Research Area: Technical/natural sciences

Publication information
Volume: 132
Issue number: 2
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Interrelation Between Sensation Level and Auditory Evoked Potentials Under Conditions of Masking Release

General Information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Epp, B. (Intern), Yasin, I. (Ekstern), Jesko, L. (Ekstern)
Publication date: 2012
Event: Abstract from 35th MidWinter Meeting of the Association for Research in Otolaryngology, San Diego, CA, United States.
Main Research Area: Technical/natural sciences
Publication: Research - peer-review › Journal article – Annual report year: 2012
Listening in adverse conditions: Masking release and effects of hearing loss

Speech perception is a complex process involving the ability to detect the speech signal, separate it from interfering sounds and decode the transmitted speech information. In contrast to normal hearing (NH) listeners, hearing-impaired (HI) listeners often show a large reduction in the masking release (MR), which is the improvement in speech intelligibility when the interferer is different from steady-state noise (e.g., a competing talker). MR is usually measured as the difference in speech reception thresholds (SRTs), the signal-to-noise ratio (SNR) where 50% of the speech is understood, and has mainly been linked to the ability to separate the target from the interferer. However, it is still not clear why HI listeners show a reduced MR and how the ability to decode the speech information is affected by impaired hearing. Thus, the purpose of this thesis was to investigate MR in both NH and HI listeners, to study the effects of hearing loss on the ability to decode speech, and to establish a framework for modeling speech intelligibility based on an auditory processing model.

The first part of the thesis established the modeling framework and showed that, by using a model that captures the processing of the different stages of the auditory system, it is possible to predict speech intelligibility using a very simple back end. Furthermore, the results indicated that the high-energy segments are the most important for speech intelligibility.

The second part focused on recent indications that the large reduction in MR observed in HI listeners is a result of measuring the MR of HI listeners at a higher signal-to-noise ratio (SNR) in stationary noise relative to NH listeners. The present work presented noise-band vocoded as well as low-pass and high-pass filtered stimuli to NH listeners, thereby decreasing their speech intelligibility and making it possible to compare the MR of NH and HI listeners not only at the same SNR, but also at the same same percent correct, which was not done in previous studies. The MR was found to be only partially related to the SRT obtained in stationary noise. Furthermore, for a competing talker, noise-vocoding strongly reduced the MR of the NH listeners to that obtained with HI listeners. This indicated that deficits in coding of temporal fine structure and fundamental frequency (F0) information may play a critical role for the reduced MR of the HI listeners.

The third part investigated the contribution of high-rate envelope fluctuations, at the output of the auditory filters, to MR. High-rate envelope fluctuations are produced by the interaction between unresolved harmonics and are related to the F0 of voiced speech. A new vocoder technique was developed to effectively attenuate the high-rate envelope fluctuations. Furthermore, high-pass filtering was used to reduce the amount of F0 information from resolved harmonics. The results showed high-rate envelope fluctuations, related to the F0, were sufficient to obtain a large MR. Furthermore, F0-related information from resolved harmonics were also sufficient for MR. However, when both high-rate envelope fluctuations and F0-related information from resolved harmonics were reduced, the MR was strongly reduced. Thus, the results indicated that F0 information is crucial for MR, but that it does not matter if it is obtained from low-order resolved harmonics or from high-rate envelope fluctuations produced by interaction between unresolved harmonics.

The final avenue of investigation focused on the effects of hearing loss on the ability to decode speech by measuring consonant confusions for both individual HI listeners and also individual utterances of the same consonants. In general, the results showed that individual HI listeners consistently confused the presented utterances with only one other consonant, and that most of the HI listeners actually made the same confusions. The results also indicated that the reason for the large variability in the confusion patterns of HI listeners observed in previous studies is that different utterances of the same consonant promote different confusions and that the HI listeners experience problems with different utterances. Overall, this thesis provides insights about the large MR observed for NH listeners, why this MR is often reduced for HI listeners and in which way impaired hearing affects the ability to decode speech information.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering
Authors: Jespersgaard, C. F. C. (Intern), Dau, T. (Intern)
Number of pages: 104
Publication date: 2012

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
ISBN (Electronic): 978-87-92465-xx-x
Original language: English
Main Research Area: Technical/natural sciences
Electronic versions: afh handling_cfc..PDF
Publication: Research › Ph.D. thesis – Annual report year: 2012

Listening talkers produce great spectral tilt contrasts
It is well known that the envelope of the long-term average speech spectrum flattens with vocal effort. A recent study [1] showed that content words had a flatter spectral envelope than content words at the same overall level for a specific Danish speech material. The present paper investigates whether this effect is present in a larger and more diverse speech material, and if the effect is greater when the talker is listening (participating in a dialogue) as compared to monologue. The monologue speech material consisted of recordings from 18 native talkers of Danish describing a network of colored geometrical shapes taken from DanPASS [2]. The spectral tilt was gauged by calculating the band-level difference in dB
between two frequency bands with pass-bands 150 to 803 Hz and 803 to 1358 Hz respectively in 5 ms intervals. This was done separately for intervals containing content words and function words and grouped by talker. The spectral tilt difference was then calculated as the average band-level difference for function words minus the average band-level difference for content words. This calculation was grouped per talker. For the monologues these differences ranged between 5 and 8 dB for the 18 talkers. Content words were defined as nouns, active verbs, adjectives and adverbs. Function words were defined as articles, pronouns, conjunctions and auxiliary verbs. Words not belonging to any of these categories were not used. The dialogue speech material was also from DanPASS and consisted of recordings from 13 of the same talkers as the monologues. In the dialogue speech aterial talkers where asked to describe a map with certain discrepancies and negotiate their way through the map. Spectral tilt differences between content- and functions words were calculated in the same way as for the monologues. The results show that the spectral tilt differences are slightly higher for dialogues than monologues. A two-way anova (grouped by talker and word type) showed that these differences are significant. We conclude that Danish talkers mark high information density in spontaneous speech (=content words) by means of flat spectral envelope, not just for monologues, but also for dialogues. Moreover, when engaged in dialogue, talkers enhance this spectral flattening. In our view it is remarkable that conclusions with statistical validity can be reached based on the over-simplified definition of spectral tilt employed in this paper. We speculate that optimizing both the definition of spectral tilt and the word categories comprising content- and function words, may allow us to observe even greater effects than reported here. The eventual goal of this line of research is to devise a simple, tractable method for distinguishing high information content from low information content in speech, based on the ubiquitous assumption that content words carry more information than function words. Such a method could potentially be applied in hearing aids, cochlear implants and automatic speech recognition.

**General information**
- State: Published
- Organisations: Department of Electrical Engineering, Hearing Systems, Copenhagen Business School
- Authors: Christiansen, T. U. (Intern), Heegård, J. (Ekstern), Henrichsen, P. J. (Ekstern)
- Publication date: 2012
- Event: Abstract from The Listening Talker, Edinburgh, United Kingdom.
- Main Research Area: Technical/natural sciences
- Spectral tilt, Spectral envelope, Speech production, Speech perception, Content words, Function words
- Electronic versions: 20120522101137202.pdf

**Metrics for performance assessment of mixed-order Ambisonics spherical microphone arrays**
Mixed-order Ambisonics (MOA) combines planar (2D) higher order Ambisonics (HOA) with lower order periphonic (3D) Ambisonics. MOA encoding from spherical microphone arrays has the potential to provide versatile recordings that can be played back using 2D, 3D or mixed systems. A procedure to generate suitable layouts for a given MOA combination order is introduced consisting of rings of microphones at several elevation angles for any given MOA combination order. Robustness and directivity measures were evaluated for four MOA layouts. Results showed that MOA vertical directivity was similar to 3D HOA and that MOA horizontal directivity was in between the planar and periphonic order.

**General information**
- State: Published
- Organisations: Department of Electrical Engineering, Hearing Systems
- Authors: Favrot, S. E. (Intern), Marschall, M. (Intern)
- Number of pages: 7
- Publication date: 2012

**Host publication information**
- Title of host publication: Proceedings of the AES 25th UK CONFERENCE
- Main Research Area: Technical/natural sciences
- Conference: AES 25th UK CONFERENCE, York, United Kingdom, 25/03/2012 - 25/03/2012
- Electronic versions: 01_favrot_aesuk2012.pdf

**Modeling auditory evoked brainstem responses to transient stimuli**
A quantitative model is presented that describes the formation of auditory brainstem responses (ABR) to tone pulses, clicks and rising chirps as a function of stimulation level. The model computes the convolution of the instantaneous discharge rates using the "humanized" nonlinear
auditory-nerve (AN) model of Zilany and Bruce (2007) and an empirically derived unitary response function which is assumed to reflect contributions from different cell populations within the auditory brainstem, recorded at a given pair of electrodes on the scalp. It is shown that the model accounts for the decrease of tone-pulse evoked wave-V latency with frequency but underestimates the level dependency of the tone-pulse as well as click-evoked latency values. Furthermore, the model correctly predicts the nonlinear wave-V amplitude behavior in response to the chirp stimulation both as a function of chirp sweeping rate and level. Overall, the results support the hypothesis that the pattern of ABR generation is strongly affected by the nonlinear and dispersive processes in the cochlea.

**General information**

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Warwick, William Demant Holding A/S
Authors: Rønne, F. M. (Intern), Dau, T. (Intern), Harte, J. (Ekstern), Elberling, C. (Ekstern)
Pages: 3903-3913
Publication date: 2012
Main Research Area: Technical/natural sciences

**Publication information**

Volume: 131
Issue number: 5
ISSN (Print): 0001-4966
Ratings:
- BFI (2018): BFI-level 2
- Web of Science (2018): Indexed yes
- BFI (2017): BFI-level 2
- Web of Science (2017): Indexed yes
- BFI (2016): BFI-level 2
- Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
- Web of Science (2016): Indexed yes
- BFI (2015): BFI-level 2
- Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
- Web of Science (2015): Indexed yes
- BFI (2014): BFI-level 2
- Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
- Web of Science (2014): Indexed yes
- BFI (2013): BFI-level 2
- Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
- ISI indexed (2013): ISI indexed yes
- Web of Science (2013): Indexed yes
- BFI (2012): BFI-level 2
- Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
- ISI indexed (2012): ISI indexed yes
- Web of Science (2012): Indexed yes
- BFI (2011): BFI-level 2
- Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
- ISI indexed (2011): ISI indexed yes
- Web of Science (2011): Indexed yes
- BFI (2010): BFI-level 2
- Scopus rating (2010): SJR 0.754 SNIP 1.528
- Web of Science (2010): Indexed yes
- BFI (2009): BFI-level 2
- Scopus rating (2009): SJR 0.783 SNIP 1.717
- Web of Science (2009): Indexed yes
Modeling speech intelligibility in adverse conditions

In everyday life, the speech we listen to is often mixed with many other sound sources as well as reverberation. In such situations, people with normal hearing are able to almost effortlessly segregate a single voice out of the background. In contrast, hearing-impaired people have great difficulty understanding speech when more than one person is talking, even when reduced audibility has been fully compensated for by a hearing aid. The reasons for these difficulties are not well understood. This presentation highlights recent concepts of the monaural and binaural signal processing strategies employed by the normal as well as impaired auditory system. Jørgensen and Dau [(2011). J. Acoust. Soc. Am. 130, 1475-1487] proposed the speech-based envelope power spectrum model (sEPSM) in an attempt to overcome the limitations of the classical speech transmission index (STI) and speech intelligibility index (SII) in conditions with nonlinearly processed speech. Instead of considering the reduction of the temporal modulation energy as the intelligibility metric, as assumed in the STI, the sEPSM applies the signal-to-noise ratio in the envelope domain (SNRenv). This metric was shown to be the key for predicting the intelligibility of reverberant speech as well as noisy speech processed by spectral subtraction. However, the sEPSM cannot account for speech subjected to phase jitter, a condition in which the spectral structure of speech is destroyed, while the broadband temporal envelope is kept largely intact. In contrast, the effects of this distortion can be predicted successfully by the spectro-temporal modulation index (STMI) [Elhilali et al., (2003). Speech Commun. 41, 331-348], which assumes an explicit analysis of the spectral modulation energy. However, since the STMI applies the same decision metric as the STI, it fails to account for spectral subtraction. The results from the different modeling approaches suggest that the SNRenv might be a key decision metric while some explicit across-frequency pre-processing seems crucial to extract relevant speech features in some conditions.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Dau, T. (Intern)
Pages: 6-6
Publication date: 2012

Host publication information
Title of host publication: The Listening Talker : Proceedings
Main Research Area: Technical/natural sciences
Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission

This paper describes the implementation and performance of a nonlinear time-domain model of the cochlea for transient stimulation and human otoacoustic emission generation. The nonlinearity simulates compressive growth of measured basilar-membrane impulse responses. The model accounts for reflection and distortion-source otoacoustic emissions (OAEs) and simulates spontaneous OAEs through manipulation of the middle-ear reflectance. The model was calibrated using human psychoacoustical and otoacoustic tuning parameters. It can be used to investigate time-dependent properties of cochlear mechanics and the generator mechanisms of otoacoustic emissions. Furthermore, the model provides a suitable preprocessor for human auditory perception models where realistic cochlear excitation patterns are desired. © 2012 Acoustical Society of America.
On the possibility of a place code for the low pitch of high-frequency complex tones

Harmonics are considered unresolved when they interact with neighboring harmonics and cannot be heard out separately. Several studies have suggested that the pitch derived from unresolved harmonics is coded via temporal fine-structure cues emerging from their peripheral interactions. Such conclusions rely on the assumption that the components of complex tones with harmonic ranks down to at least 9 were indeed unresolved. The present study tested this assumption via three different measures: (1) the effects of relative component phase on pitch matches, (2) the effects of dichotic presentation on pitch matches, and (3) listeners’ ability to hear out the individual components. No effects of relative component phase or dichotic presentation on pitch matches were found in the tested conditions. Large individual differences were found in listeners’ ability to hear out individual components. Overall, the results are consistent with the coding of individual harmonic frequencies, based on the tonotopic activity pattern or phase locking to individual harmonics, rather than with temporal coding of single-channel interactions. However, they are also consistent with more general temporal theories of pitch involving the across-channel summation of information from resolved and/or unresolved harmonics. Simulations of auditory-nerve responses to the stimuli suggest potential benefits to a spatiotemporal mechanism.
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 0.848 SNIP 1.633
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 0.865 SNIP 1.647
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 0.752 SNIP 1.559
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 0.954 SNIP 1.749
Web of Science (2005): Indexed yes
Scopus rating (2004): SJR 0.77 SNIP 1.787
Web of Science (2004): Indexed yes
Scopus rating (2003): SJR 0.882 SNIP 1.712
Web of Science (2003): Indexed yes
Scopus rating (2002): SJR 0.87 SNIP 1.501
Web of Science (2002): Indexed yes
Scopus rating (2001): SJR 0.719 SNIP 1.467
Web of Science (2001): Indexed yes
Scopus rating (2000): SJR 0.621 SNIP 1.411
Web of Science (2000): Indexed yes
Scopus rating (1999): SJR 0.591 SNIP 1.319
Perceptual Confusions Among Consonants, Revisited: Cross-Spectral Integration of Phonetic-Feature Information and Consonant Recognition

The perceptual basis of consonant recognition was experimentally investigated through a study of how information associated with phonetic features (Voicing, Manner, and Place of Articulation) combines across the acoustic-frequency spectrum. The speech signals, 11 Danish consonants embedded in Consonant + Vowel + Liquid syllables, were partitioned into 3/4-octave bands (“slits”) centered at 750 Hz, 1500 Hz, and 3000 Hz, and presented individually and in two- or three-slit combinations. The amount of information transmitted (IT) was calculated from consonant-confusion matrices for each feature and slit combination. The growth of IT was measured as a function of the number of slits presented and their center frequency for the phonetic features and consonants. The IT associated with Voicing, Manner, and Consonants sums nearly linearly for two-band stimuli irrespective of their center frequency. Adding a third band increases the IT by an amount somewhat less than predicted by linear cross-spectral integration (i.e., a compressive function). In contrast, for Place of Articulation, the IT gained through addition of a second or third slit is far more than predicted by linear, cross-spectral summation. This difference is mirrored in a measure of error-pattern similarity across bands—Symmetric Redundancy. Consonants, as well as Voicing and Manner, share a moderate degree of redundancy between bands. In contrast, the cross-spectral redundancy associated with Place is close to zero, which means the bands are essentially independent in terms of decoding this feature. Because consonant recognition and Place decoding are highly correlated (correlation coefficient r² = 0.99), these results imply that the auditory processes underlying consonant recognition are not strictly linear. This may account for why conventional cross-spectral integration speech models, such as the Articulation Index, Speech Intelligibility Index, and the Speech Transmission Index do not predict intelligibility and segment recognition well under certain conditions (e.g., discontiguous frequency bands, audio-visual speech).
Prediction of speech masking release for fluctuating interferers based on the envelope power signal-to-noise ratio

The speech-based envelope power spectrum model (sEPSM) presented by Jørgensen and Dau [(2011). J. Acoust. Soc. Am. 130, 1475-1487] estimates the envelope signal-to-noise ratio (SNRenv) after modulation-frequency selective processing. This approach accurately predicts the speech intelligibility for normal-hearing listeners in conditions with additive stationary noise, reverberation, and nonlinear processing with spectral subtraction. The latter condition represents a case in which the standardized speech intelligibility index and the speech transmission index fail. However, the sEPSM is limited to conditions with stationary interferers due to the long-term estimation of the envelope power and cannot account for the well-known phenomenon of speech masking release. Here, a short-term version of the sEPSM is described [Jørgensen and Dau, 2012, in preparation], which estimates the SNRenv in short temporal segments. Predictions obtained with the short-term sEPSM are compared to data from Kjems et al. [(2009). J. Acoust. Soc. Am. 126 (3), 1415-1426] where speech is mixed with four different interferers, including speech-shaped noise, bottle noise, car noise, and a highly non-stationary cafe noise. The model accounts well for the differences in intelligibility observed for the stationary and non-stationary interferers, demonstrating further that the SNRenv is crucial for speech comprehension.
the envelope SNR in 10-ms time frames. Predictions obtained with the short-term sEPSM are compared to data from Kjems et al. [2009]. J. Acoust. Soc. Am. 126 (3), 1415-1426] where speech is mixed with four different interferers, including speech-shaped noise, bottle noise, car noise, and a highly non-stationary cafe noise. The model accounts well for the differences in intelligibility observed for the stationary and non-stationary interferers, demonstrating further that the envelope SNR is crucial for speech comprehension.
Objective: To establish reference hearing threshold levels for chirps and frequency-specific chirps. Design: Hearing thresholds were determined monaurally for broad-band chirps and octave-band chirps using the Etymotic Research, ER-3A insert earphone. The chirps were presented using two repetition rates, 20 and 90 stimuli/s, and with alternating polarity in blocks of one second duration. The test procedure and test conditions were in accordance with the recommendations given in ISO 389-9 (2009). The ascending method (ISO 8253-1, 2010) was applied using a step size of 5 dB. The chirps were played back from a Tucker Davies Technologies System II, and a Matlab program controlled the test setup. The results are specified in dB peak-to-peak equivalent threshold sound pressure levels (dB peETSPL). Study sample: The test group consisted of 25 otologically-normal young adults (age 18–25 years). Results: The results are in good agreement with the results from another investigation of hearing thresholds using the same chirp stimuli, and the values for the octave-band chirps are in line with the standardized reference values for corresponding tone bursts (ISO 389-6, 2007). Conclusions: The results of the present investigation are relevant for the international standard on short duration signals, ISO 389-6 (2007).
Relating binaural pitch perception to the individual listener's auditory profile

The ability of eight normal-hearing listeners and fourteen listeners with sensorineural hearing loss to detect and identify pitch contours was measured for binaural-pitch stimuli and salience-matched monaurally detectable pitches. In an effort to determine whether impaired binaural pitch perception was linked to a specific deficit, the auditory profiles of the individual listeners were characterized using measures of loudness perception, cognitive ability, binaural processing, temporal fine structure processing, and frequency selectivity, in addition to common audiometric measures. Two of the listeners were found not to perceive binaural pitch at all, despite a clear detection of monaural pitch. While both binaural and monaural pitches were detectable by all other listeners, identification scores were significantly lower for binaural than for monaural pitch. A total absence of binaural pitch sensation coexisted with a loss of a binaural signal-detection advantage in noise, without implying reduced cognitive function. Auditory filter bandwidths did not correlate with the difference in pitch identification scores between binaural and monaural pitches. However, subjects with impaired binaural pitch perception showed deficits in temporal fine structure processing. Whether the observed deficits stemmed from peripheral or central mechanisms could not be resolved here, but the present findings may be useful for hearing loss characterization. (C) 2012 Acoustical Society of America. [http://dx.doi.org/10.1121/1.3689554]
Relationship between masking release in fluctuating maskers and speech reception thresholds in stationary noise

In contrast to normal-hearing (NH) listeners, hearing-impaired (HI) listeners often show strongly reduced masking release (MR) in fluctuating interferers, which has commonly been associated with spectral and temporal processing deficits. However, it has recently been proposed that the reduced MR could result from an increased speech recognition threshold (SRT) in stationary noise [Bernstein and Grant, J. Acoust. Soc. Am. 125, 3358-3372 (2009)]. This was tested by presenting noise-band vocoded as well as low-pass and high-pass filtered stimuli to NH listeners, thereby increasing their stationary-noise SRTs to those of the HI listeners. If the primary determinant of MR is the SRT in stationary noise then the amount of the MR should be independent of the type of processing used to obtain the stationary-noise SRT. However, the relation between the amount of MR and the stationary-noise SRT depended on the type of processing. For a fluctuating interferer, none of the processing conditions reduced the MR of the NH listeners to that of the HI listeners. In contrast, for an interfering talker, the results for vocoded stimuli were similar to those of the HI listeners. Overall, these results suggest that the observed MR is only partially related to the stationary-noise SRT.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Centre for Applied Hearing Research
Authors: Christiansen, C. (Intern), Dau, T. (Intern)
Pages: 1655-1666
Publication date: 2012
Main Research Area: Technical/natural sciences

Publication information
Volume: 132
Issue number: 3
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
Relationships between Cochlear Tuning and Delay Probed with a Nonlinear Transmission-Line Model

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Epp, B. (Intern), Bergevin, C. (Ekstern), Shera, C. (Ekstern)
Publication date: 2012
Representational Similarity Analysis Reveals Heterogeneous Networks Supporting Speech Motor Control

The everyday act of speaking involves the complex processes of speech motor control. One important feature of such control is regulation of articulation when auditory concomitants of speech do not correspond to the intended motor gesture. While theoretical accounts of speech monitoring posit multiple functional components required for detection of errors in speech planning (e.g., Levelt, 1983), neuroimaging studies generally indicate either single brain regions sensitive to speech production errors, or small, discrete networks. Here we demonstrate that the complex system controlling speech is supported by a complex neural network that is involved in linguistic, motoric and sensory processing. With the aid of novel real-time acoustic analyses and representational similarity analyses of fMRI signals, our data show functionally differentiated networks underlying auditory feedback control of speech.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Queen's University, Medical Research Council
Authors: Zheng, Z. (Ekstern), Cusack, R. (Ekstern), Johnsrude, I. (Ekstern), MacDonald, E. (Intern), Munhall, K. (Ekstern), Vicente-Grabovetsky, A. (Ekstern)
Number of pages: 3
Publication date: 2012
Event: Poster session presented at 18th Annual Meeting of the Organization for Human Brain Mapping, Beijing, China.
Main Research Area: Technical/natural sciences
Source: dtu
Source-ID: u::4306
Publication: Research - peer-review › Poster – Annual report year: 2012

Robustness of a Mixed-Order Ambisonics Microphone Array for Sound Field Reproduction

Spherical microphone arrays can be used to capture and reproduce the spatial characteristics of acoustic scenes. A mixed-order Ambisonics (MOA) approach was recently proposed to improve the horizontal spatial resolution of microphone arrays with a given number of transducers. In this paper, the performance and robustness of an MOA array to variations in microphone characteristics as well as self-noise was investigated. Two array processing strategies were evaluated. Results showed that the expected performance benefits of MOA are achieved at mid to high frequencies, and that robustness to various errors was similar to that of HOA arrays with both strategies. The approach based on minimizing the error of the reproduced spherical harmonic functions showed better performance at high frequencies for the MOA layout.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, National Acoustic Laboratories
Authors: Marschall, M. (Intern), Favrot, S. E. (Intern), Buchholz, J. (Ekstern)
Number of pages: 11
Publication date: 2012
Host publication Information
Title of host publication: 132nd Audio Engineering Society convention
Publisher: AES
Main Research Area: Technical/natural sciences
Conference: AES 132nd Convention, Budapest, Hungary, 26/04/2012 - 26/04/2012
Source: dtu
Source-ID: u::4309
Publication: Research - peer-review › Article in proceedings – Annual report year: 2012

Sense Meets Nonsense: a dual-layer Danish speech corpus for perception studies

In this paper, we present the newly established Danish speech corpus PiTu. The corpus consists of recordings of 28 native Danish talkers (14 female and 14 male) each reproducing (i) a series of nonsense syllables, and (ii) a set of authentic natural language sentences. The speech corpus is tailored for investigating the relationship between early stages of the speech perceptual process and later stages. We present our considerations involved in preparing the experimental set-up, producing the anechoic recordings, compiling the data, and exploring the materials in linguistic research. We report on a small pilot experiment demonstrating how PiTu and similar speech corpora can be used in studies of prosody as a function of semantic content. The experiment addresses the issue of whether the governing principles of Danish prosody assignment is mainly talker-specific or mainly content-typical (under the specific experimental
Word Recognition for Temporally and Spectrally Distorted Materials: The Effects of Age and Hearing Loss

Objectives: The purpose of Experiment 1 was to measure word recognition in younger adults with normal hearing when speech or babble was temporally or spectrally distorted. In Experiment 2, older listeners with near-normal hearing and with hearing loss (for pure tones) were tested to evaluate their susceptibility to changes in speech level and distortion types. The results across groups and listening conditions were compared to assess the extent to which the effects of the distortions on word recognition resembled the effects of age-related differences in auditory processing or pure-tone conditions.
hearing loss.

Design: In Experiment 1, word recognition was measured in 16 younger adults with normal hearing using Northwestern University Auditory Test No. 6 words in quiet and the Words-in-Noise test distorted by temporal jittering, spectral smearing, or combined jittering and smearing. Another 16 younger adults were evaluated in four conditions using the Words-in-Noise test in combinations of unaltered or jittered speech and unaltered or jittered babble. In Experiment 2, word recognition in quiet and in babble was measured in 72 older adults with near-normal hearing and 72 older adults with hearing loss in four conditions: unaltered, jittered, smeared, and combined jittering and smearing.

Results: For the listeners in Experiment 1, word recognition was poorer in the distorted conditions compared with the unaltered condition. The signal to noise ratio at 50% correct word recognition was 4.6 dB for the unaltered condition, 6.3 dB for the jittered, 6.8 dB for the smeared, 6.9 dB for the double-jitter, and 8.2 dB for the combined jitter-smear conditions. Jittering both the babble and speech signals did not significantly reduce performance compared with jittering only the speech. In Experiment 2, the older listeners with near-normal hearing and hearing loss performed best in the unaltered condition, followed by the jitter and smear conditions, with the poorest performance in the combined jitter-smear condition in both quiet and noise. Overall, listeners with near-normal hearing performed better than listeners with hearing loss by similar to 30% in quiet and similar to 6 dB in noise. In the quiet distorted conditions, when the level of the speech was increased, performance improved for the hearing loss group, but decreased for the older group with near-normal hearing. Recognition performance of younger listeners in the jitter-smear condition and the performance of older listeners with near-normal hearing in the unaltered conditions were similar. Likewise, the performance of older listeners with near-normal hearing in the jitter-smear condition and the performance of older listeners with hearing loss in the unaltered conditions were similar.

Conclusions: The present experiments advance our understanding regarding how spectral or temporal distortions of the fine structure of speech affect word recognition in older listeners with and without clinically significant hearing loss. The Speech Intelligibility Index was able to predict group differences, but not the effects of distortion. Individual differences in performance were similar across all distortion conditions with both age and hearing loss being implicated. The speech materials needed to be both spectrally and temporally distorted to mimic the effects of age-related differences in auditory processing and hearing loss.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, East Tennessee State University, Linköping University
Authors: Smith, S. L. (Ekstern), Pichora-Fuller, M. K. (Ekstern), Wilson, R. H. (Ekstern), MacDonald, E. (Intern)
Pages: 349-366
Publication date: 2012
Main Research Area: Technical/natural sciences

Publication Information
Journal: Ear and Hearing
Volume: 33
Issue number: 3
ISSN (Print): 0196-0202
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 2.97 SJR 1.865 SNIP 1.571
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 1.753 SNIP 2.008 CiteScore 2.94
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 1.93 SNIP 1.726 CiteScore 2.86
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 1.893 SNIP 2.109 CiteScore 3.18
ISI indexed (2013): ISI indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 1.918 SNIP 1.708 CiteScore 2.95
Processing of spatial sounds in the impaired auditory system

Understanding speech in complex acoustic environments presents a challenge for most hearing-impaired listeners. In conditions where normal-hearing listeners effortlessly utilize spatial cues to improve speech intelligibility, hearing-impaired listeners often struggle. In this thesis, the influence of two such cues on speech intelligibility was studied. First, the benefit from early reflections (ER's) in a room was determined using a virtual auditory environment. ER’s were found to be useful for speech intelligibility, but to a smaller extent than the direct sound (DS). The benefit was quantified with an intelligibility-weighted “efficiency factor” which revealed that the spectral characteristics of the ER’s caused the reduced benefit. Hearing-impaired listeners were able to utilize the ER energy as effectively as normal-hearing listeners, most likely because binaural processing was not required for the integration of the ER’s with the DS. Different masker types were found to have an impact on the binaural processing of the overall speech signal but not on the processing of ER’s. Second, the influence of interaural level differences (ILD’s) on speech intelligibility was investigated with a hearing aid research platform. ILD’s are considered important for localizing sounds and for the perceptual separation of competing sound sources. Bilateral hearing aids with independent compression algorithms typically decrease ILD’s, such that the perception of spatial sounds becomes distorted. Hearing aids that are binaurally linked can utilize the signals at both ears and preserve the ILD’s through co-ordinated compression. Hearing-impaired listeners received a small, but not significant advantage from linked compared to independent compression. It was concluded that, for speech intelligibility, the exact ILD information is not crucial. The results from an additional experiment demonstrated that the ER benefit was maintained with independent as well as with linked hearing aid compression. Overall, this work contributes to the understanding of ER processing in listeners with normal and impaired hearing and may have implications for speech perception models and the development of compensation strategies in future generations of hearing instruments.

General information

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Arweiler, I. (Intern), Buchholz, J. (Intern), Dau, T. (Intern)
Number of pages: 117
Publication date: Sep 2011

Place of publication: Kgs. Lyngby, Denmark
A cross-language study of compensation in response to real-time formant perturbation

Past studies have shown that when formants are perturbed in real time, speakers spontaneously compensate for the perturbation by changing their formant frequencies in the opposite direction to the perturbation. Further, the pattern of these results suggests that the processing of auditory feedback error operates at a purely acoustic level. This hypothesis was tested by comparing the response of three language groups to real-time formant perturbations, (1) native English speakers producing an English vowel /e/, (2) native Japanese speakers producing a Japanese vowel (=e=), and (3) native Japanese speakers learning English, producing /e/. All three groups showed similar production patterns when F1 was decreased; however, when F1 was increased, the Japanese groups did not compensate as much as the native English speakers. Due to this asymmetry, the hypothesis that the compensatory production for formant perturbation operates at a purely acoustic level was rejected. Rather, some level of phonological processing influences the feedback processing behavior.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Queen's University, University of Western Ontario
Authors: Mitsuya, T. (Ekstern), MacDonald, E. (Intern), Purcell, D. W. (Ekstern), Munhall, K. G. (Ekstern)
Pages: 2978-2986
Publication date: 2011
Main Research Area: Technical/natural sciences
Application of a circular 2D hard-sphere microphone array for higher-order Ambisonics auralization

A circular microphone array mounted on a rigid sphere was realized and its application to higher-order Ambisonics (HOA) auralization was analysed. Besides the 2D Ambisonics application this array design provides a promising basis for the development of a mixed-order Ambisonics recording system, in which the high spatial resolution of a high-order 2D system is efficiently combined with a periphonic sound field representation of a low-order 3D system. In order to evaluate and optimize the performance of the spherical 2D microphone array, a versatile simulation framework was developed, which included the entire processing path from the sound source via the (noisy) microphone signals and the loudspeaker output signals to the reproduced sound field in the centre of a loudspeaker array. This framework was then used to analyse the response of the different system stages to an ideal plane wave. The simulation results showed very good agreement with corresponding plane wave recordings in an anechoic chamber and thus, confirming the general applicability of the simulation framework. An overall preference listening test was performed to estimate the optimal array radius and amount of regularization, two (dependent) parameters that mainly determine the balance between low frequency directionality, signal coloration and microphone noise amplification. The different stimuli were created with the framework using different values for both the array radius and the regularization coefficient lambda. It was shown that best results were achieved with a radius of around 5-10 cm and a lambda of about 0.01.

General information

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Weller, T. (Ekstern), Favrot, S. E. (Intern), Buchholz, J. (Intern)
Pages: 2269-2274
Are binaural recordings needed for subjective and objective annoyance assessment of traffic noise?
Humans are annoyed when they are exposed to environmental noise. Traditional measures such as sound pressure levels may not correlate well with how humans perceive annoyance, therefore it is important to investigate psychoacoustic metrics that may correlate better with the perceived annoyance of environmental noise than the A-weighted equivalent sound pressure level. This study examined whether the use of binaural recordings of sound events improves the correlation between the objective metrics and the perceived annoyance, particularly for road traffic noise. Metrics based on measurement with a single microphone and on binaural sound field recordings have been examined and compared. In order to acquire data for the subjective perception of annoyance, a series of listening tests has been carried out. It is concluded that binaural loudness metrics from binaural recordings are better correlated with the subjective annoyance assessment.

Behavioral estimates of basilar-membrane input-output in normal-hearing listeners
To characterize human cochlear processing it would be beneficial to behaviorally estimate the basilar membrane (BM) input-output (I/O) function. In recent studies, forward masking has been used to estimate BM compression. In this study, a growth-of-forward-masking (GOM) paradigm (e.g., Oxenham and Plack, 1997) was extended to also estimate the knee point of the I/O function between linear and compressive processing. If a low-level signal is masked by an on-frequency masker, such that the signal is processed linearly and the masker compressively according to the I/O function, then a steeper GOM function is expected than that obtained for a high-level signal where both masker and signal are processed compressively. The knee point can be estimated at the input level where the GOM slope changes significantly. Data were collected from seven normal - hearing listeners. The method was found to provide estimates of the BM I/O function for a wider range of input levels than in previously suggested methods, due to the additional estimates of the knee points.
Can a static nonlinearity account for the dynamics of otoacoustic emission suppression?

This study investigates whether time-dependent compression mechanisms in the cochlea are necessary to explain dynamic properties of otoacoustic emissions (OAEs). Dynamic properties of click-evoked OAEs (CEOAEs) have been observed in temporal suppression; the effect where the CEOAE magnitude is reduced when a click is presented less than 10ms before the test click. A timedomain model of the cochlea that represented the basilar membrane (BM) as a cascade of coupled bandpass filters was used to investigate the cochlear origin of temporal suppression in CEOAEs. The model, implemented with a time-invariant nonlinearity, was able to simulate temporal suppression, but was unable to account for the exact time scale and magnitude of the effect. The results suggest that temporal overlap of BM impulse responses can account for suppression in CEOAEs, but that an additional time-dependent cochlear gain mechanism may be needed to account the high suppression maxima at inter-click intervals larger than zero.

Characterizing auditory processing and perception in individual listeners with sensorineural hearing loss

This study considered consequences of sensorineural hearing loss in ten listeners. The characterization of individual hearing loss was based on psychoacoustic data addressing audiometric pure-tone sensitivity, cochlear compression, frequency selectivity, temporal resolution, and intensity discrimination. In the experiments it was found that listeners with comparable audiograms can show very different results in the supra-threshold measures. In an attempt to account for the observed individual data, a model of auditory signal processing and perception [Jepsen et al., J. Acoust. Soc. Am. 124, 422–438 (2008)] was used as a framework. The parameters of the cochlear processing stage of the model were adjusted to account for behaviorally estimated individual basilar-membrane input/output functions and the audiogram, from which the amounts of inner hair-cell and outer hair-cell losses were estimated as a function of frequency. All other model parameters were left unchanged. The predictions showed a reasonably good agreement with the measured individual data in the frequency selectivity and forward masking conditions while the variation of intensity discrimination thresholds across listeners was underestimated by the model. The model and the associated parameters for individual hearing-impaired listeners might be useful for investigating effects of individual hearing impairment in more complex conditions, such as speech intelligibility in noise.
<table>
<thead>
<tr>
<th>Publication information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume: 129</td>
</tr>
<tr>
<td>Issue number: 1</td>
</tr>
<tr>
<td>ISSN (Print): 0001-4966</td>
</tr>
<tr>
<td>Ratings:</td>
</tr>
<tr>
<td>BFI (2018): BFI-level 2</td>
</tr>
<tr>
<td>Web of Science (2018): Indexed yes</td>
</tr>
<tr>
<td>BFI (2017): BFI-level 2</td>
</tr>
<tr>
<td>Web of Science (2017): Indexed yes</td>
</tr>
<tr>
<td>BFI (2016): BFI-level 2</td>
</tr>
<tr>
<td>Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27</td>
</tr>
<tr>
<td>Web of Science (2016): Indexed yes</td>
</tr>
<tr>
<td>BFI (2015): BFI-level 2</td>
</tr>
<tr>
<td>Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77</td>
</tr>
<tr>
<td>Web of Science (2015): Indexed yes</td>
</tr>
<tr>
<td>BFI (2014): BFI-level 2</td>
</tr>
<tr>
<td>Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8</td>
</tr>
<tr>
<td>Web of Science (2014): Indexed yes</td>
</tr>
<tr>
<td>BFI (2013): BFI-level 2</td>
</tr>
<tr>
<td>Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2</td>
</tr>
<tr>
<td>ISI indexed (2013): ISI indexed yes</td>
</tr>
<tr>
<td>Web of Science (2013): Indexed yes</td>
</tr>
<tr>
<td>BFI (2012): BFI-level 2</td>
</tr>
<tr>
<td>Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75</td>
</tr>
<tr>
<td>ISI indexed (2012): ISI indexed yes</td>
</tr>
<tr>
<td>Web of Science (2012): Indexed yes</td>
</tr>
<tr>
<td>BFI (2011): BFI-level 2</td>
</tr>
<tr>
<td>Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68</td>
</tr>
<tr>
<td>ISI indexed (2011): ISI indexed yes</td>
</tr>
<tr>
<td>Web of Science (2011): Indexed yes</td>
</tr>
<tr>
<td>BFI (2010): BFI-level 2</td>
</tr>
<tr>
<td>Scopus rating (2010): SJR 0.754 SNIP 1.528</td>
</tr>
<tr>
<td>Web of Science (2010): Indexed yes</td>
</tr>
<tr>
<td>BFI (2009): BFI-level 2</td>
</tr>
<tr>
<td>Scopus rating (2009): SJR 0.783 SNIP 1.717</td>
</tr>
<tr>
<td>Web of Science (2009): Indexed yes</td>
</tr>
<tr>
<td>BFI (2008): BFI-level 2</td>
</tr>
<tr>
<td>Scopus rating (2008): SJR 0.848 SNIP 1.633</td>
</tr>
<tr>
<td>Web of Science (2008): Indexed yes</td>
</tr>
<tr>
<td>Scopus rating (2007): SJR 0.865 SNIP 1.647</td>
</tr>
<tr>
<td>Web of Science (2007): Indexed yes</td>
</tr>
<tr>
<td>Scopus rating (2006): SJR 0.752 SNIP 1.559</td>
</tr>
<tr>
<td>Web of Science (2006): Indexed yes</td>
</tr>
<tr>
<td>Scopus rating (2005): SJR 0.954 SNIP 1.749</td>
</tr>
<tr>
<td>Web of Science (2005): Indexed yes</td>
</tr>
<tr>
<td>Scopus rating (2004): SJR 0.77 SNIP 1.787</td>
</tr>
<tr>
<td>Web of Science (2004): Indexed yes</td>
</tr>
</tbody>
</table>
Normal-hearing (NH) listeners can typically better understand speech in the presence of a fluctuating noise or a competing talker compared to a stationary noise interferer. However, for hearing-impaired (HI) listeners, this masking release (MR) is strongly reduced or completely absent. Traditionally, this has been attributed to the ability of NH listeners to utilize the speech in the low-amplitude periods of the masker, an ability that is supposed to be reduced for HI listeners due to reduced temporal and spectral resolution. However, [1] proposed that the reduced MR experienced by HI listeners is due to their higher speech reception threshold (SRT) in stationary noise. In the present study, this hypothesis was investigated by presenting noise-band vocoded as well as low-pass and high-pass filtered stimuli to the NH listeners. In this way, the SRTs of the NH listeners were similar to those of the HI listeners in stationary noise. It is shown that the relation between the MR and the SRT in stationary noise clearly depends on the type of processing used. The results therefore demonstrate that the SRT in stationary noise has only a minor effect on the amount of MR, which is in contrast to [1].

Planar (2D) and periphonic (3D) higher-order Ambisonics (HOA) playback systems are widely used in multi-channel audio applications. For a given Ambisonics order, 2D systems require far less loudspeakers and provide a larger spatial resolution but cannot naturally reproduce elevated sound sources. In order to combine the benefits of 2D and 3D systems, a higher order 2D playback system can be mixed with a lower order 3D system. In the present study, a mixed-order Ambisonics playback system was realised by extending the spherical harmonics decomposition of a 3D sound field with additional horizontal components. The performance of the system was analysed by considering a small and a large loudspeaker setup, allowing for different combinations of 2D and 3D Ambisonics orders. An objective evaluation showed that the systems provided a high spatial resolution for horizontal sources while producing a smooth decrease in spatial resolution with increasing source elevation until 3D performance is reached. This observation was confirmed by a listening test (simulated concert scenario), which showed that in comparison to a conventional 3D system the perceived spatial resolution for sources in the horizontal plane can be significantly increased by adding 2D components and thereby approaching 2D system's performance. Simultaneously, frequency spectrum properties of horizontal sound sources were
Two objective measures of human cochlear tuning, using stimulus-frequency otoacoustic emissions (SFOAE), have been proposed. One measure used SFOAE phase-gradient delay and the other twotone suppression (2TS) tuning curves. Here, it is hypothesized that the two measures lead to different frequency functions in the same listener. Two experiments were conducted in ten young adult normal-hearing listeners in three frequency bands (1-2 kHz, 3-4 kHz and 5-6 kHz). Experiment 1 recorded SFOAE latency as a function of stimulus frequency, and experiment 2 recorded 2TS isoinput tuning curves. In both cases, the output was converted into a sharpness-of-tuning factor based on the equivalent rectangular bandwidth. In both experiments, sharpness-of-tuning curves were shown to be frequency dependent, yielding sharper relative tuning with increasing frequency. Only a weak frequency dependence of the sharpness-of-tuning curves was observed for experiment 2, consistent with objective and behavioural estimates from the literature. Most importantly, the absolute difference between the two tuning estimates was very large and statistically significant. It is argued that the 2TS estimates of cochlear tuning likely represents the underlying properties of the suppression mechanism, and not necessarily cochlear tuning. Thus the phase-gradient delay estimate is the most likely one to reflect cochlear tuning.
Investigating the periodicity of transient-evoked otoacoustic emission envelopes

This study investigates the cochlear origin of the multiple temporal lobes that are often observed in the transient-evoked otoacoustic emission (TEOAE) envelope. This "waxing and waning" of the OAE amplitude can be observed in tone-burst (TB) OAEs and sometimes also in click-evoked (CE) OAEs. TBOAE envelopes were analyzed for several frequency and level configurations to investigate the relation between the envelope periodicity and the characteristic frequency of the emission component. It was found that the TBOAE envelope periodicity was dominated by the modulation caused by the interaction of several CEOAE and SOAE frequency components close to the stimulus frequency. A second and smaller TBOAE envelope component showed a qualitative agreement with envelope periods predicted by a model based on multiple cochlear reflections. Multiple reflections between the frequency components in the TBOAE and the middle-ear boundary may contribute to the TBOAE envelope periodicity, but were not the main modulation component in waxing and waning of the investigated TBOAEs.
Low-frequency versus high-frequency synchronisation in chirp-evoked auditory brainstem responses

This study investigates the frequency specific contribution to the auditory brainstem response (ABR) of chirp stimuli. Frequency rising chirps were designed to compensate for the cochlear traveling wave delay, and lead to larger wave-V amplitudes than for click stimuli as more auditory nerve fibres fire synchronously. Traditional click stimuli were believed to only excite high-frequency fibres synchronously. It is still currently unclear whether the broad-band chirp stimulus leads to increased synchronisation of both low- and high-frequency fibres. It is also unclear if both these groups of fibres contribute significantly to the overall wave-V amplitude. In the present study, ABRs were recorded from 10 normal-hearing listeners using low- and high-frequency band-limited chirps and clicks (0.1 – 1.5 kHz and 1.5 - 10 kHz) presented at a level of 40 dB HL. The results showed significantly larger wave-V amplitudes for both low and high-frequency band-limited chirps than for the filtered clicks. This demonstrates that the synchronisation of nerve fibres occurs across the entire frequency range at this presentation level, and this leads to significant increases in wave-V amplitudes. The increase for the low-frequency chirp was found to be clearly larger than that obtained at the higher frequencies.
Modeling auditory grouping based on a temporal coherence analysis

Current models of auditory streaming rely primarily on frequency separation for sound segregation. However, spectral components that are well separated in frequency are no longer heard as separate streams if presented synchronously rather than consecutively. Elhilali et al. [1] suggested a conceptual model to account for grouping based on synchrony, but the model was not evaluated with experimental data. In this study, it was experimentally tested how the temporal overlap between spectral components affected the perception of one or two streams. This was investigated for a range of tone-repetition rates. The data suggested that the perceptual organization depends on the absolute asynchrony of the tones. When the asynchrony of the tones was less than 20-44 ms, the tones were organized into the same perceptual stream, and when the asynchrony was larger, two separate streams were perceived. The conceptual model performed well in the tested conditions, however, some issues associated with the peripheral processing stages in the model were observed, caused by the frequency dependent delay of the filters. An alternative peripheral model is thus required for the conceptual model to function properly.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering
Authors: Christiansen, S. K. (Intern), Jepsen, M. L. (Intern), Dau, T. (Intern)
Pages: 2673-2678
Publication date: 2011

Host publication information
Title of host publication: Proceedings of Forum Acusticum 2011
ISBN (Print): 978-84-694-1520-7
Main Research Area: Technical/natural sciences
Links: http://www.fa2011.org/
Source: orbit
Source-ID: 278709
Publication: Research - peer-review › Article in proceedings – Annual report year: 2011

Modeling the effects of compression and suppression on estimates of auditory frequency selectivity

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, National Acoustic Laboratories
Authors: Marschall, M. (Intern), Buchholz, J. (Ekstern), Dau, T. (Intern)
Pages: 605-606
Publication date: 2011

Host publication information
Title of host publication: Fortschritte der Akustik : DAGA 2011 Deutsche Gesellschaft für Akustik
ISBN (Print): 978-3-939296-02-7
Main Research Area: Technical/natural sciences
Conference: Fortschritte der Akustik - DAGA 2011, 01/01/2011
Source: orbit
Source-ID: 276986
Publication: Research - peer-review › Article in proceedings – Annual report year: 2011

Modelling the level-dependent latency of the auditory brainstem response

Auditory brainstem responses (ABR) are used for both clinical and research purposes to objectively assess human hearing. A prominent feature of the transient evoked ABR is the level-dependent latency of the distinct peaks in its waveform. The latency of the most prominent peak, wave-V, is about 8 ms at a peak equivalent sound pressure level of 55
dB, and reduces for increasing level by approximately 1 ms / 20 dB. A classical explanation for this finding asserts that an increasing stimulus levels lead to a broadened excitation pattern on the basilar membrane. This results in further activation of the basal regions of the cochlea. Given the physical properties of the basilar membrane, increased basal activation is believed to cause a decreasing ABR latency. An Auditory Nerve (AN) model and the Dual Resonance Non-Linearity (DRNL) filter model are considered as separate front-end cochlear models to simulate ABRs. Even though both models incorporate level-dependent tuning and synapse adaptation, and thus theoretically should be capable of simulating level-dependent latencies, both models under-predict the latencies. The failure to produce accurate simulations suggests, that the level-depending tuning in the models is not accurately modelled. The level dependency of the basilar membrane filter tuning in humans is not well described in the literature and could therefore cause the modelling difficulties.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, University of Warwick, William Demant Holding A/S
Authors: Rønne, F. M. (Intern), Dau, T. (Intern), Harte, J. (Ekstern), Elberling, C. (Ekstern)
Publication date: 2011

Host publication information
Title of host publication: Proceedings of Forum Acusticum 2011
ISBN (Print): 978-84-694-1520-7
Main Research Area: Technical/natural sciences
Links:
http://www.fa2011.org/
Source: orbit
Source-ID: 278708
Publication: Research - peer-review › Article in proceedings – Annual report year: 2011

Modulation filtering using an optimization approach to spectrogram reconstruction
Modulations across time and frequency are known from previous studies to play a significant role for speech intelligibility. Hence, well-chosen manipulations of modulations via an accurate tool to systematically modify the modulation content of a signal might be useful for the improvement of speech intelligibility. This study investigates modulation filtering in a time-frequency representation of the signal (e.g., a spectrogram), using a novel approach for reconstructing a signal from its modified representation. It is suggested that this synthesis is regarded as an optimization problem, where the variables are the time samples of the output signal and where the cost function to minimize is the difference between the target spectrogram and the current spectrogram. This approach is made feasible, with regard to the large number of variables involved, by use of a limited-memory optimization algorithm. This study presents basic results regarding temporal modulation filtering and discusses the novel method and its possibilities of improvement.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering
Authors: Decorsiere, R. J. B. (Intern), Søndergaard, P. L. (Intern), Buchholz, J. (Intern), Dau, T. (Intern)
Publication date: 2011

Host publication information
Title of host publication: Proceedings of Forum Acusticum 2011
Publisher: European Acoustics Association
ISBN (Print): 978-84-694-1520-7
Main Research Area: Technical/natural sciences
Electronic versions:
Decorsiere2011[1].pdf
Source: orbit
Source-ID: 282053
Publication: Research - peer-review › Article in proceedings – Annual report year: 2011

Neural coding and perception of pitch in the normal and impaired human auditory system
Pitch is an important attribute of hearing that allows us to perceive the musical quality of sounds. Besides music perception, pitch contributes to speech communication, auditory grouping, and perceptual segregation of sound sources. In this work, several aspects of pitch perception in humans were investigated using psychophysical methods. First, hearing loss was found to affect the perception of binaural pitch, a pitch sensation created by the binaural interaction of noise stimuli. Specifically, listeners without binaural pitch sensation showed signs of retrocochlear disorders. Despite adverse effects of reduced frequency selectivity on binaural pitch perception, the ability to accurately process the temporal fine structure (TFS) of sounds at the output of the cochlear filters was found to be essential for perceiving binaural pitch.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering
Authors: Decorsiere, R. J. B. (Intern), Søndergaard, P. L. (Intern), Buchholz, J. (Intern), Dau, T. (Intern)
Publication date: 2011

Host publication information
Title of host publication: Proceedings of Forum Acusticum 2011
Publisher: European Acoustics Association
ISBN (Print): 978-84-694-1520-7
Main Research Area: Technical/natural sciences
Electronic versions:
Decorsiere2011[1].pdf
Source: orbit
Source-ID: 282053
Publication: Research - peer-review › Article in proceedings – Annual report year: 2011
Monaural TFS processing also played a major and independent role for a variety of basic auditory tasks, indicating that it may be a crucial measure to consider for hearing-loss characterization. In contrast to hearing-impaired listeners, adults with dyslexia showed no deficits in binaural pitch perception, suggesting intact low-level auditory mechanisms. The second part of this work investigated the role of temporal and spectral information for complex pitch perception. In particular, it was shown that the low pitch evoked by high-frequency complex tones was not conveyed by temporal envelope cues as such. Moreover, the fact that the individual frequency components could not be heard out separately by the listeners suggested that the low pitch relied on TFS information, even in high-frequency regions where phase-locking in auditory nerve cells is believed to be weak. A second set of experiments could however not validate the assumption of a temporally-coded pitch and indicated that the use of spectral cues remained plausible. Simulations of auditory-nerve representations of the complex tones further suggested that a spectrotimetrical mechanism combining precise timing information across auditory channels might best account for the behavioral data. Overall, this work provides insights into the fundamental auditory mechanisms underlying pitch perception, and may have implications for future pitch-perception models, as well as strategies for auditory-profile characterization and restoration of accurate pitch perception in impaired hearing.

**On the relationship between multi-channel envelope and temporal fine structure**

The envelope of a signal is broadly defined as the slow changes in time of the signal, where as the temporal fine structure (TFS) are the fast changes in time, i.e. the carrier wave(s) of the signal. The focus of this paper is on envelope and TFS in multi-channel systems. We discuss the difference between a linear and a non-linear model of information-extraction from the envelope, and show that using a non-linear method for information-extraction, it is possible to obtain almost all information about the originating signal. This is shown mathematically and numerically for different kinds of systems providing an increasingly better approximation to the auditory system. A corollary from these results is that it is not possible to generate a test signal containing contradictory information in its multi-channel envelope and TFS.
Perception of interaural time differences at threshold and suprathreshold levels

Experiments were performed to study and compare the perception of interaural time differences (ITDs) at threshold and suprathreshold levels. In a first experiment, ITD thresholds were measured for 1-ERB-wide stimuli centered at frequencies between 150 and 1500 Hz. The interaural correlation was systematically varied between 1 and 0.85. For fully correlated signals, the dependence of the thresholds on the center frequency was in line with previous results, showing the highest ITD sensitivity between 750 and 1000 Hz. Thresholds measured for partially correlated stimuli were elevated but showed a similar dependence on frequency as the thresholds measured for fully correlated stimuli. In a second experiment, the task of the listeners was to align a broadband pointer signal carrying a single ITD, on a broadband target signal that carried a different ITD in each frequency band. The ITDs in the target signals were either close to threshold or at suprathreshold levels. The results showed that, in contrast to the dominance region between 750 and 1000 Hz observed in the case of close-to-thresholds presentation, ITD information is equally integrated across frequency for suprathreshold values.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering
Authors: Le Goff, N. (Intern), Buchholz, J. (Intern), Dau, T. (Intern)
Publication date: 2011

Perceptual Weights for Loudness Judgments of 6-Tone Complexes

In a series of studies, 6 subjects with normal hearing (NH) and 3 with sensorineural hearing loss (SNHL) judged the overall loudness of 6-tone complexes comprised of octave frequencies from 0.25 to 8 kHz. In two tasks, tones were equated in level in dB SPL or in sensation level (SL) and a range of SPL or SL values was tested. Both tasks used two-interval forced-choice trials, with level “jitter” introduced by selecting the level of each tone from a normal distribution with specified mean level and standard deviation of 5 dB. Subjects were instructed to indicate which complex was louder. In the “loudness” task there was no difference in mean level across the two intervals. In the “sample discrimination” task, the two complexes differed by an average of 5 dB. For both tasks, perceptual weights were derived by correlating the differences in level between matched-frequency tones in the complexes and the loudness decision on each trial. Weights derived from the loudness task (no mean level difference) were highly correlated with weights derived from the sample discrimination task (5-dB difference). For SPL conditions, both NH and SNHL subjects placed less weight on the lowest frequency and greater weight on higher frequencies with increasing intensity of the complexes, with larger effects for NH subjects. This effect was not observed in conditions where levels were equated in SL. Weights derived from a single-interval categorical loudness scaling task, where subjects judged the overall loudness of 6-tone complexes, were highly correlated with those from the other tasks. Simulation of these experiments using a model of loudness perception [Moore and Glasberg, J. Hear Res. 188 (70-88)] yielded weights for these stimuli that were highly correlated with specific loudness, but the observed weights did not agree with predicted weights. This suggests that model assumptions regarding specific loudness are not correct.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Boys Town National Research Hospital
Authors: Jesteadt, W. (Ekstern), Valente, D. L. (Ekstern), Joshi, S. N. (Intern)
Publication date: 2011
Event: Abstract from Thirty-Fourth Annual Midwinter Research Meeting of the Association for Research in Otolaryngology, San Diego, Californien, United States.
Main Research Area: Technical/natural sciences
Electronic versions:
Preview of “2011; Perceptual Weights...nts of 6-Tone Complexes”.pdf
Source: dtu
Source-ID: u::7802
Publication: Research - peer-review › Conference abstract for conference – Annual report year: 2011

Perceptual weights for loudness reflect central spectral processing.
Weighting patterns for loudness obtained using the reverse correlation method are thought to reveal the relative contributions of different frequency regions to total loudness, the equivalent of specific loudness. Current models of
loudness assume that specific loudness is determined by peripheral processes such as compression and masking. Here we test this hypothesis using 20-tone harmonic complexes (200Hz f0, 200 to 4000Hz, 250 ms, 65 dB/Component) added in opposite phase relationships (Schroeder positive and negative). Due to the varying degree of envelope modulations, these time-reversed harmonic complexes have been shown to produce different outputs at the basilar membrane and different amounts of forward and simultaneous masking. The perceptual weights for loudness did not differ for these two complexes. To determine whether the level rove introduced to obtain weights had changed the fundamental differences in the stimuli, a similar level rove (68 dB) was introduced on each component of Schroeder positive and negative forward maskers. The Schroder negative maskers continued to be more effective. These results suggest that perceptual weights for loudness are not completely determined by peripheral processes and reflect a central frequency weighting template.
Pitch perception beyond the traditional existence region of pitch

Humans’ ability to recognize musical melodies is generally limited to pure-tone frequencies below 4 or 5 kHz. This limit coincides with the highest notes on modern musical instruments and is widely believed to reflect the upper limit of precise stimulus-driven spike timing in the auditory nerve. We tested the upper limits of pitch and melody perception in humans using pure and harmonic complex tones, such as those produced by the human voice and musical instruments, in melody recognition and pitch-matching tasks. We found that robust pitch perception can be elicited by harmonic complex tones with fundamental frequencies below 2 kHz, even when all of the individual harmonics are above 6 kHz—well above the currently accepted existence region of pitch and above the currently accepted limits of neural phase locking. The results suggest that the perception of musical pitch at high frequencies is not constrained by temporal phase locking in the auditory nerve but may instead stem from higher-level constraints shaped by prior exposure to harmonic sounds.

General information

State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, University of Minnesota
Authors: Oxenham, A. J. (Ekstern), Micheyl, C. (Ekstern), Keebler, M. V. (Ekstern), Loper, A. (Ekstern), Santurette, S. (Intern)
Pages: 7629-7634
Publication date: 2011
Main Research Area: Technical/natural sciences

Publication information

Journal: Proceedings of the National Academy of Sciences of the United States of America
Volume: 108
Issue number: 18
ISSN (Print): 0027-8424
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 8.56 SJR 6.321 SNIP 2.629
Web of Science (2016): Indexed yes
Predicting speech intelligibility based on the signal-to-noise envelope power ratio after modulation-frequency selective processing

A model for predicting the intelligibility of processed noisy speech is proposed. The speech-based envelope power spectrum model has a similar structure as the model of Ewert and Dau [(2000). J. Acoust. Soc. Am. 108, 1181-1196],
developed to account for modulation detection and masking data. The model estimates the speech-to-noise envelope power ratio, SNR env, at the output of a modulation filterbank and relates this metric to speech intelligibility using the concept of an ideal observer. Predictions were compared to data on the intelligibility of speech presented in stationary speech-shaped noise. The model was further tested in conditions with noisy speech subjected to reverberation and spectral subtraction. Good agreement between predictions and data was found in all cases. For spectral subtraction, an analysis of the model's internal representation of the stimuli revealed that the predicted decrease of intelligibility was caused by the estimated noise envelope power exceeding that of the speech. The classical concept of the speech transmission index fails in this condition. The results strongly suggest that the signal-to-noise ratio at the output of a modulation frequency selective process provides a key measure of speech intelligibility. © 2011 Acoustical Society of America.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering
Authors: Jørgensen, S. (Intern), Dau, T. (Intern)
Pages: 1475-1487
Publication date: 2011
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of the Acoustical Society of America
Volume: 130
Issue number: 3
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 0.848 SNIP 1.633
Predicting speech intelligibility in adverse conditions: evaluation of the speech-based envelope power spectrum model

The speech-based envelope power spectrum model (sEPSM) [Jørgensen and Dau (2011). J. Acoust. Soc. Am., 130 (3), 1475–1487] estimates the envelope signal-to-noise ratio (SNRenv) of distorted speech and accurately describes the speech recognition thresholds (SRT) for normal-hearing listeners in conditions with additive noise, reverberation, and nonlinear processing by spectral subtraction. The latter represents a condition where the standardized speech intelligibility index and speech transmission index fail. However, the sEPSM is limited to stationary interferers due to the fact that predictions are based on the long-term SNRenv. As an attempt to extend the model to deal with fluctuating interferers, a short-time version of the sEPSM is presented. The SNRenv of a speech sample is estimated from a combination of SNRenv-values calculated in short time frames. The model is evaluated in adverse conditions by comparing predictions to measured data from [Kjems et al. (2009). J. Acoust. Soc. Am. 126 (3), 1415-1426] where speech is mixed with four different interferers, including speech-shaped noise, bottle noise, car noise, and cafe noise. The model accounts well for the differences in intelligibility observed for the different interferers. None of the standardized models successfully describe these data.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Jørgensen, S. (Intern), Dau, T. (Intern)
Pages: 307-314
Publication date: 2011

Host publication information
Title of host publication: Proceedings of the 3rd International Symposium on Auditory and Audiological Research: Speech Perception and Auditory Disorders
Publisher: The Danavox Jubilee Foundation
Editors: Dau, T., Jepsen, M. L., Poulsen, T., Dalsgaard, J. C.
ISBN (Print): 978-87-990013-3-0
Main Research Area: Technical/natural sciences
Predicting the effect of spectral subtraction on the speech recognition threshold based on the signal-to-noise ratio in the envelope domain

Digital noise reduction strategies are important in technical devices such as hearing aids and mobile phones. One well-described noise reduction scheme is the spectral subtraction algorithm. Many versions of the spectral subtraction scheme have been presented in the literature, but the methods have rarely been evaluated perceptually in terms of speech intelligibility. This study analyzed the effects of the spectral subtraction strategy proposed by Berouti et al. [ICASSP 4 (1979), 208-211] on the speech recognition threshold (SRT) obtained with sentences presented in stationary speech-shaped noise. The SRT was measured in five normal-hearing listeners in six conditions of spectral subtraction. The results showed an increase of the SRT after processing, i.e., a decreased speech intelligibility, in contrast to what is predicted by the Speech Transmission Index (STI). Here, another approach is proposed, denoted the speech-based envelope power spectrum model (sEPSM) which predicts the intelligibility based on the signal-to-noise ratio in the envelope domain. In contrast to the STI, the sEPSM is sensitive to the increased amount of the noise envelope power as a consequence of the spectral subtraction operation, which leads to a decreased speech intelligibility in this model, in quantitative agreement with the experimental data.

Predicting the intelligibility of processed noisy speech based on the signal-to-noise ratio in the modulation domain

This study investigates behavioural and objective measures of temporal auditory processing and their relation to the ability to understand speech in noise. The experiments were carried out on a homogeneous group of seven hearing-impaired listeners with normal sensitivity at low frequencies (up to 1 kHz) and steeply sloping hearing losses above 1 kHz. For comparison, data were also collected for five normal-hearing listeners. Temporal processing was addressed at low frequencies by means of psychoacoustical frequency discrimination, binaural masked detection and amplitude modulation (AM) detection. In addition, auditory brainstem responses (ABRs) to clicks and broadband rising chirps were recorded. Furthermore, speech reception thresholds (SRTs) were determined for Danish sentences in speech-shaped noise. The main findings were: (1) SRTs were neither correlated with hearing sensitivity as reflected in the audiogram nor with the AM detection thresholds which represent an envelope-based measure of temporal resolution; (2) SRTs were correlated with frequency discrimination and binaural masked detection which are associated with temporal fine-structure coding; (3) The wave-V thresholds for the chirp-evoked ABRs indicated a relation to SRTs and the ability to process temporal fine structure. Overall, the results demonstrate the importance of low-frequency temporal processing for speech reception.
which can be affected even if pure-tone sensitivity is close to normal.
Sound Exposure of Symphony Orchestra Musicians

Background: Assessment of sound exposure by noise dosimetry can be challenging especially when measuring the exposure of classical orchestra musicians where sound originate from many different instruments. A new measurement method of bilateral sound exposure of classical musicians was developed and used to characterize sound exposure of the left and right ear simultaneously in two different symphony orchestras. Objectives: To measure binaural sound exposure of professional classical musicians and to identify possible exposure risk factors of specific musicians. Methods: Sound exposure was measured with microphones mounted on the musician’s ears and recorded digitally. The recorded sound was analysed and the specific sound exposure of the left and the right ear was determined for the musicians. A total of 114 measurements covering 106 h were recorded in two symphony orchestras. Results: Sound exposure depends significantly on the specific instrument and the repertoire played by the exposed musician. Concerts, group rehearsals and individual practice were all significant contributors to the sound exposure. The highest LAeq of 86 –98 dB was found among the brass players. High string players were exposed from 82 to 98 dBA and their left ear was exposed 4.6 dB more than the right ear. Percussionists were exposed to high sound peaks >115 dBC but less continuous sound exposure was observed in this group. Musicians were exposed up to LAeq8h of 92 dB and a majority of musicians were exposed to sound levels exceeding LAeq8h of 85 dB. Conclusions: Binaural recording of the individual sound exposure showed that orchestra musicians could be exposed differently to the left and right ear and that they were primarily exposed from their own instruments. Specific repertoires as well as the specific instrument determine the level of exposure.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Odense University Hospital, University of Southern Denmark
Authors: Schmidt, J. H. (Ekstern), Pedersen, E. R. (Ekstern), Juhl, P. M. (Ekstern), Christensen-Dalsgaard, J. (Ekstern), Andersen, T. D. (Ekstern), Poulsen, T. (Intern), Bælum, J. (Ekstern)
Pages: 893-905
Publication date: 2011
Main Research Area: Technical/natural sciences

Publication information
Journal: Annals of Work Exposures and Health
Volume: 55
Issue number: 8
ISSN (Print): 2398-7308
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 1
Scopus rating (2016): SJR 0.789 SNIP 1.048 CiteScore 1.44
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 0.873 SNIP 1.445 CiteScore 1.92
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 1.047 SNIP 1.556 CiteScore 1.81
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 0.987 SNIP 1.343 CiteScore 1.91
ISI indexed (2013): ISI indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): SJR 0.925 SNIP 1.44 CiteScore 1.88
Speech recognition employing biologically plausible receptive fields

The main idea of the project is to build a widely speaker-independent, biologically motivated automatic speech recognition (ASR) system. The two main differences between our approach and current state-of-the-art ASRs are that i) the features used here are based on the responses of neuronlike spectro-temporal receptive fields to auditory spectrogram input, motivated by the auditory pathway of humans, and ii) the adaptation or learning algorithms involved are biologically inspired. This is in contrast to state-of-the-art combinations of Mel-frequency cepstral coefficients and Hidden Markov Model-based adaptation procedures. Two databases are used, TI46 for discrete speech a subset of the TIMIT database collected from speakers belonging to the New York dialect region. Each of the selection of 10 sentences is uttered once by each of 35 speakers. The major differences between the two data sets initiate the development and comparison of two distinct ASRs within the project, which will be presented in the following. Employing a reduced sampling frequency and bandwidth of the signals, the ASR algorithm reaches and goes beyond recognition results that are known from humans.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Centre for Applied Hearing Research
Authors: Fereczkowski, M. (Intern), Bothe, H. (Intern)
Pages: 73-78
Publication date: 2011

Host publication information
Title of host publication: Proceedings of Forum Acusticum
Publisher: European Acoustics Association
Main Research Area: Technical/natural sciences
Hidden Markov models, Learning algorithms, Speech recognition
Source: dtu
Source-ID: n:oai:DTIC-ART:compendex/370501977::31775
Publication: Research - peer-review › Article in proceedings – Annual report year: 2011
Temporal and Spectral Cues for Musical Timbre Perception in Electric Hearing

The purpose of this study was to investigate musical timbre perception in cochlear-implant (CI) listeners using a multidimensional scaling technique to derive a timbre space. Methods: Sixteen stimuli that synthesized western musical instruments were used (McAdams, Winsberg, Donnadieu, De Soete, & Krimphoff, 1995). Eight CI listeners and 15 normal hearing (NH) listeners participated. Each listener made judgments of dissimilarity between stimulus pairs. Acoustical analyses that characterized the temporal and spectral characteristics of each stimulus were performed to examine the psychophysical nature of each perceptual dimension. For NH listeners, the timbre space was best represented in three dimensions, one correlated with the temporal envelope (log-attack time) of the stimuli, one correlated with the spectral envelope (spectral centroid), and one correlated with the spectral fine structure (spectral irregularity) of the stimuli. The timbre space from CI listeners, however, was best represented by two dimensions, one correlated with temporal envelope features and the other weakly correlated with spectral envelope features of the stimuli. Temporal envelope was a dominant cue for timbre perception in CI listeners. Compared to NH listeners, CI listeners showed reduced reliance on both spectral envelope and spectral fine structure cues for timbre perception.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Northeastern University
Authors: Kong, Y. (Ekstern), Mullangi, A. (Ekstern), Marozeau, J. (Intern), Epstein, M. (Ekstern)
Pages: 981-995
Publication date: 2011
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of Speech, Language, and Hearing Research
Volume: 54
Issue number: 3
ISSN (Print): 1092-4388
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 2.18 SJR 1.152 SNIP 1.418
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 0.985 SNIP 1.153 CiteScore 1.88
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 1.244 SNIP 1.686 CiteScore 2.45
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 1.275 SNIP 1.774 CiteScore 2.44
ISI indexed (2013): ISI indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): SJR 1.229 SNIP 1.708 CiteScore 2.32
ISI indexed (2012): ISI indexed yes
BFI (2011): BFI-level 1
Scopus rating (2011): SJR 1.297 SNIP 1.725 CiteScore 2.35
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 1
Scopus rating (2010): SJR 1.632 SNIP 1.785
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 1
Scopus rating (2009): SJR 1.503 SNIP 1.743
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 1.177 SNIP 1.417
Scopus rating (2007): SJR 1.277 SNIP 1.521
Scopus rating (2006): SJR 1.249 SNIP 1.474
Scopus rating (2005): SJR 1.17 SNIP 1.504
Temporal integration of loudness measured using categorical loudness scaling and matching procedures

Temporal integration of loudness of 1 kHz tones with 5 and 200 ms durations was assessed in four subjects using two loudness measurement procedures: categorical loudness scaling (CLS) and loudness matching. CLS provides a reliable and efficient procedure for collecting data on the temporal integration of loudness and previously reported nonmonotonic behavior observed at mid-sound pressure level levels is replicated with this procedure. Stimuli that are assigned to the same category are effectively matched in loudness, allowing the measurement of temporal integration with CLS without curve-fitting, interpolation, or assumptions concerning the form of the loudness growth function.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Boys Town National Research Hospital
Authors: Valente, D. L. (Ekstern), Joshi, S. N. (Intern), Jesteadt, W. (Ekstern)
Pages: EL32-EL37
Publication date: 2011
Main Research Area: Technical/natural sciences

Publication information
Journal: J A S A Express Letters
Volume: 130
Issue number: 1
ISSN (Print): 1529-7853
Ratings:
BFI (2018): BFI-level 1
BFI (2017): BFI-level 1
BFI (2016): BFI-level 1
BFI (2015): BFI-level 1
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
ISI indexed (2013): ISI indexed yes
BFI (2012): BFI-level 1
ISI indexed (2012): ISI indexed yes
BFI (2011): BFI-level 1
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 1
BFI (2009): BFI-level 1
BFI (2008): BFI-level 1
Scopus rating (2008): SJR 0.403 SNIP 0.81
Scopus rating (2007): SJR 0.544 SNIP 0.775
Scopus rating (2006): SJR 0.354 SNIP 0.599
Scopus rating (2005): SJR 0.6 SNIP 0.794
Temporal suppression of the click-evoked otoacoustic emission level-curve

The click-evoked otoacoustic emission (CEOAE) level-curve grows linearly for clicks below 40–60 dB and saturates for higher inputs. This study investigates dynamic (i.e., time-dependent) features of the CEOAE level-curve by presenting a suppressor-click less than 8 ms before the test-click. An alteration of the CEOAE level-curve, designated here as temporal suppression, was observed within this time period, and was shown to depend on the levels and the temporal separation of the two clicks. Temporal suppression occurred for all four subjects tested, and resulted in a vertical offset from the unsuppressed level-curve for test-click levels greater than 50 dB peak-equivalent level (peSPL). Temporal suppression was greatest for suppressors presented 1–4 ms before the test click, and the magnitude and time scale of the effect were subject dependent. Temporal suppression was furthermore observed for the short- (i.e., 6–18 ms) and long-latency (i.e., 24–36 ms) regions of the CEOAE, indicating that temporal suppression similarly affects synchronized spontaneous otoacoustic emissions (SSOAEs) and purely evoked CEOAE components. Overall, this study demonstrates that temporal suppression of the CEOAE level-curve reflects a dynamic process in human cochlear processing that works on a time scale of 0–10 ms.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering
Authors: Verhulst, S. (Intern), Harte, J. (Intern), Dau, T. (Intern)
Pages: 1452-1463
Publication date: 2011
Main Research Area: Technical/natural sciences

Publication information
Volume: 129
Issue number: 3
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
The influence of spectral characteristics of early reflections on speech intelligibility
The auditory system takes advantage of early reflections (ERs) in a room by integrating them with the direct sound (DS) and thereby increasing the effective speech level. In the present paper the benefit from realistic ERs on speech intelligibility in diffuse speech-shaped noise was investigated for normal-hearing and hearing-impaired listeners. Monaural and binaural speech intelligibility tests were performed in a virtual auditory environment where the spectral characteristics of ERs from a simulated room could be preserved. The useful ER energy was derived from the speech intelligibility results and the efficiency of the ERs was determined as the ratio of the useful ER energy to the total ER energy. Even though ER energy contributed to speech intelligibility, DS energy was always more efficient, leading to better speech intelligibility for both groups of listeners. The efficiency loss for the ERs was mainly ascribed to their altered spectrum compared to the DS and to the filtering by the torso, head, and pinna. No binaural processing other than a binaural summation effect could be observed.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering
Authors: Arweiler, I. (Intern), Buchholz, J. (Intern)
Pages: 996-1005
Publication date: 2011
Main Research Area: Technical/natural sciences

Publication information
The Linear Time Frequency Analysis Toolbox

The Linear Time Frequency Analysis Toolbox is a Matlab/Octave toolbox for computational time-frequency analysis. It is intended both as an educational and computational tool. The toolbox provides the basic Gabor, Wilson and MDCT transform along with routines for constructing windows (filter prototypes) and routines for manipulating coefficients. It also provides a bunch of demo scripts devoted either to demonstrating the main functions of the toolbox, or to exemplify their use in specific signal processing applications. In this paper we describe the used algorithms, their mathematical background as well as some signal processing applications.

General information

State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Université de Provence, Austrian Academy of Sciences
Authors: Søndergaard, P. L. (Intern), Torrésani, B. (Ekstern), Balazs, P. (Ekstern)
Pages: 1250032
Publication date: 2011
Main Research Area: Technical/natural sciences

Publication information

Journal: International Journal of Wavelets, Multiresolution and Information Processing
Volume: 10
Issue number: 4
ISSN (Print): 0219-6913
Ratings:

Web of Science (2018): Indexed yes
Web of Science (2017): Indexed Yes
Scopus rating (2016): SJR 0.266 SNIP 0.392 CiteScore 0.51
Scopus rating (2015): SJR 0.311 SNIP 0.437 CiteScore 0.52
Scopus rating (2014): SJR 0.342 SNIP 0.812 CiteScore 0.68
Scopus rating (2013): SJR 0.367 SNIP 0.881 CiteScore 0.95
ISI indexed (2013): ISI indexed yes
Scopus rating (2012): SJR 0.426 SNIP 1.086 CiteScore 1.45
ISI indexed (2012): ISI indexed yes
Scopus rating (2011): SJR 0.516 SNIP 1.276 CiteScore 1.3
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
Scopus rating (2010): SJR 0.513 SNIP 1.043
Scopus rating (2009): SJR 0.272 SNIP 0.423
Scopus rating (2008): SJR 0.149 SNIP 0.39
Scopus rating (2007): SJR 0.157 SNIP 0.12
Original language: English
Numerical harmonic analysis, Time-frequency transforms, Gabor transform, Signal processing, Representation of systems, Efficient algorithms, Mathematical software

DOIs:
10.1142/S0219691312500324
Source: orbit
Source-ID: 282998
Publication: Research - peer-review › Journal article – Annual report year: 2012

The role of temporal fine structure information for the low pitch of high-frequency complex tones

The fused low pitch evoked by complex tones containing only unresolved high-frequency components demonstrates the ability of the human auditory system to extract pitch using a temporal mechanism in the absence of spectral cues. However, the temporal features used by such a mechanism have been a matter of debate. For stimuli with components
lying exclusively in high-frequency spectral regions, the slowly varying temporal envelope of sounds is often assumed to
be the only information contained in auditory temporal representations, and it has remained controversial to what extent
the fast amplitude fluctuations, or temporal fine structure (TFS), of the conveyed signal can be processed. Using a pitch-
matching paradigm, the present study found that the low pitch of inharmonic transposed tones with unresolved
components was consistent with the timing between the most prominent TFS maxima in their waveforms, rather than
envelope maxima. Moreover, envelope cues did not take over as the absolute frequency or rank of the lowest component
was raised and TFS cues thus became less effective. Instead, the low pitch became less salient. This suggests that
complex pitch perception does not rely on envelope coding as such, and that TFS representation might persist at higher
frequencies than previously thought.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Santurette, S. (Intern), Dau, T. (Intern)
Pages: 282-292
Publication date: 2011
Main Research Area: Technical/natural sciences

Publication information
Volume: 129
Issue number: 1
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 0.848 SNIP 1.633
Web of Science (2008): Indexed yes
**Towards a binaural modelling toolbox**

The Auditory Modelling Toolbox (AMToolbox) is a new Matlab/Octave toolbox for developing and applying auditory perceptual models and in particular binaural models. The philosophy behind the project is that the models should be implemented in a consistent manner, well documented and user-friendly in order to allow students and researchers to actively work with current models and further develop existing ones. In addition to providing the models, it is a goal of the project to collect published human data and definitions of model experiments. This will simplify the verification of models by running the model experiments and comparing the predictions to human data. The software is released under the GNU Public License (GPL) version 3, and can be downloaded from http://amtoolbox.sourceforge.net.

**General information**

**State**: Published  
**Organisations**: Hearing Systems, Department of Electrical Engineering, Cardiff University, Technische Universität Berlin, Austrian Academy of Sciences  
**Publication date**: 2011

**Host publication information**

**Title of host publication**: Proceedings of Forum Acusticum  
**ISBN (Print)**: 978-84-694-1520-7  
**Main Research Area**: Technical/natural sciences  
**Conference**: Forum Acusticum 2011, Aalborg, Denmark, 26/06/2011 - 26/06/2011  
**Source**: orbit  
**Source-ID**: 282660

**Publication**: Research - peer-review › Article in proceedings – Annual report year: 2011

**A loudspeaker-based room auralization system for auditory research**

In complex acoustic environments, such as a train station or a café, hearing-impaired people often experience difficulties to communicate even when wearing hearing instruments, whereas normal-hearing people are typically able to communicate without effort in such conditions. In order to systematically study the signal processing of realistic sounds by normal-hearing and hearing-impaired listeners, a flexible, reproducible and fully controllable auditory environment is needed. A loudspeaker-based room auralization (LoRA) system was developed in this thesis to provide virtual auditory environments (VAEs) with an array of loudspeakers. The LoRA system combines state-of-the-art acoustic room models with sound-field reproduction techniques. Limitations of these two techniques were taken into consideration together with the limitations of the human auditory system to localize sounds in reverberant environments. Each part of the early
incoming sound to the listener was auralized with either higher-order Ambisonic (HOA) or using a single loudspeaker. The late incoming sound was auralized with a specific algorithm in order to provide a diffuse reverberation with minimal coloration artifacts. In order to assess the usability of the LoRA system, one objective and two subjective evaluations were carried out. The objective evaluation showed that the physical characteristics of the acoustic scenario were preserved by the involved signal processing of the system. The first subjective evaluation assessed the impact of the auralization technique used for the early incoming sound (HOA or single loudspeaker) on speech intelligibility. A listening test showed that speech intelligibility experiments can be reliably conducted with the LoRA system with both techniques. The second evaluation investigated the perception of distance in VAEs generated by the LoRA system. These results showed that the distance of far field sources are similarly perceived in these VAEs as in real environments. For close sources (<1 m), a comprehensive study about the near field compensated HOA method was presented and an alternative post-processing was proposed that allowed for the perception of very close sound sources, nearly as accurately as real sources. Beside investigating the auditory system, such virtual auditory environments (VAEs) are also relevant for evaluating and optimizing hearing instruments and communication devices.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Favrot, S. E. (Intern), Buchholz, J. (Intern), Dau, T. (Intern)
Number of pages: 117
Publication date: Jun 2010

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
ISBN (Print): 978-87-92465-23-8
Original language: English

Series: CONTRIBUTIONS TO HEARING RESEARCH
Number: 9
Main Research Area: Technical/natural sciences
Loudspeaker-based auralization, Speech intelligibility, Virtual auditory environment, LoRA, Acoustic, Distance discrimination
Electronic versions:
SFPPhdThesis_Final0.pdf
Source: orbit
Source-ID: 262226
Publication: Research › Ph.D. thesis – Annual report year: 2010

Modeling auditory processing and speech perception in hearing-impaired listeners
A better understanding of how the human auditory system represents and analyzes sounds and how hearing impairment affects such processing is of great interest for researchers in the fields of auditory neuroscience, audiology, and speech communication as well as for applications in hearing-instrument and speech technology. In this thesis, the primary focus was on the development and evaluation of a computational model of human auditory signal-processing and perception. The model was initially designed to simulate the normal-hearing auditory system with particular focus on the nonlinear processing in the inner ear, or cochlea. The model was shown to account for various aspects of spectro-temporal processing and perception in tasks of intensity discrimination, tone-in-noise detection, forward masking, spectral masking and amplitude modulation detection. Secondly, a series of experiments was performed aimed at experimentally characterizing the effects of cochlear damage on listeners' auditory processing, in terms of sensitivity loss and reduced temporal and spectral resolution. The results showed that listeners with comparable audiograms can have very different estimated cochlear input-output functions, frequency selectivity, intensity discrimination limens and effects of simultaneous- and forward masking. Part of the measured data was used to adjust the parameters of the stages in the model, that simulate the cochlear processing. The remaining data were used to evaluate the fitted models. It was shown that an accurate simulation of cochlear input-output functions, in addition to the audiogram, played a major role in accounting both for sensitivity and supra-threshold processing. Finally, the model was used as a front-end in a framework developed to predict consonant discrimination in a diagnostic rhyme test. The framework was constructed such that discrimination errors originating from the front-end and the back-end were separated. The front-end was fitted to individual listeners with cochlear hearing loss according to non-speech data, and speech data were obtained in the same listeners. It was shown that most observations in the measured consonant discrimination error patterns were predicted by the model, although error rates were systematically underestimated by the model in few particular acoustic-phonetic features. These results reflect a relation between basic auditory processing deficits and reduced speech perception performance in the listeners with cochlear hearing loss. Overall, this work suggests a possible explanation of the variability in consequences of cochlear hearing loss. The proposed model might be an interesting tool for, e.g., evaluation of hearing-aid signal processing.

General information
State: Published
A loudspeaker-based room auralisation (LoRA) system for auditory perception research

Most research on understanding the signal processing of the auditory system has been realized in anechoic or almost anechoic environments. The knowledge derived from these experiments cannot be directly transferred to reverberant environments. In order to investigate the auditory signal processing of reverberant sounds, a loudspeaker-based room auralisation (LoRA) system is proposed here. The LoRA system efficiently combines modern room acoustic modelling techniques with loudspeaker-based auralization (i.e., single loudspeakers, higher-order Ambisonics). Thereby, aspects of the auditory precedence effect are utilized to realise highly authentic room reverberation. This system aims at providing a flexible research platform for conducting auditory experiments with normal-hearing, hearing-impaired, and aided hearing-impaired listeners in a fully controlled and realistic environment. An overall description of the LoRA processing is first presented, followed by a battery of objective and subjective tests to demonstrate the applicability of the different components of the system. In the objective evaluation, monaural and binaural room acoustic measures (e.g., reverberation time, clarity, interaural cross correlation coefficient) were considered. The subject evaluation included speech intelligibility and distance perception measures.

Behavioral Measures of Monaural Temporal Fine Structure Processing

Deficits in temporal fine structure (TFS) processing found in hearing-impaired listeners have been shown to correlate poorly to audibility and frequency selectivity, despite adverse effects on speech perception in noise. This underlines the need for an independent measure of TFS processing when characterizing hearing impairment. Estimating the acuity of
monaural TFS processing in humans however remains a challenge. One suggested measure is based on the ability of listeners to detect a pitch shift between harmonic (H) and inharmonic (I) complex tones with unresolved components (e.g. Moore et al., JASA 125:3412-3422, 2009). However, spectral cues arising from detectable excitation pattern shifts or audible combination tones might supplement TFS cues in this H/I-discrimination task. The present study further assessed the importance of the role of TFS, in contrast to that of temporal envelope and spectral resolution, for the low pitch evoked by high-frequency complex tones. The aim was to estimate the efficiency of monaural TFS cues as a function of the stimulus center frequency $F_c$ and its ratio $N$ to the stimulus envelope repetition rate. A pitch-matching paradigm was used, such that changes in spectral indices were not useable as a cue. The low pitch of broadband pulse-trains was matched to that of inharmonic transposed tones ($3sFcs\leq 7$ kHz, $N=[11.5,14.5]$). Resolvability of the stimulus components was assessed, and the contribution of TFS information to individual pitch-matching results was estimated and compared to performance of the same subjects in an H/I-discrimination experiment ($2.2sFcs9$ kHz, $N=[11,13,15]$) similar to that of Moore et al. (2009). Pitch matches revealed an ambiguous low pitch related to the timing between TFS peaks near adjacent envelope maxima, up to $F_c=7$ kHz for $N=11.5$ and $F_c=5$ kHz for $N=14.5$. Pitch salience decreased as $F_c$ or $N$ increased, but pitch matches never relied on the envelope repetition rate. Moreover, the results from the component-resolvability experiment indicated an inability of subjects to hear out the lowest frequency components of the stimuli. This strongly suggests that the monaural representation of TFS persists at high frequencies and prevails over envelope or spectral cues for perception of the low pitch of high-frequency complex tones. Individual performance in the H/I-discrimination experiment is therefore expected to show a similar dependency on $F_c$ and $N$ as the corresponding individual pitch matching results. If so, such methods may be useful to estimate the upper frequency limit for monaural TFS processing in individual subjects.

**General information**
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Santurette, S. (Intern), Dau, T. (Intern)
Publication date: 2010
Event: Poster session presented at International Hearing Aid Research Conference, Lake Tahoe, CA, United States.
Main Research Area: Technical/natural sciences
Links: [http://www.hei.org/ihcon/](http://www.hei.org/ihcon/)
Source: orbit
Source-ID: 267025
Publication: Research - peer-review › Poster – Annual report year: 2010

**Binaural Processing and Spatial Hearing**

**General information**
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Number of pages: 534
Publication date: 2010

**Publication Information**
Publisher: The Danavox Jubilee Foundation
ISBN (Print): 87-990013-2-2
Original language: English
Series: Proceedings of ISAAF
Main Research Area: Technical/natural sciences
Links: [http://www.isaar.eu/previous.html](http://www.isaar.eu/previous.html)
Source: orbit
Source-ID: 274976
Publication: Research - peer-review › Book – Annual report year: 2010

**Characterizing and modeling dynamic processes in the cochlea using otoacoustic emissions**

An important characteristic of human hearing is that it amplifies weak sounds while attenuating louder ones. This gain transformation takes place in the inner ear (i.e., cochlea), and is responsible for a compressive relation between the level of the presented and perceived sound. The cochlear gain mechanism is essential for our hearing and degrades when hearing impairment develops. A comprehensive understanding of the gain involved in the intact human cochlea is crucial, as hearing instruments try to compensate for the loss in cochlear gain caused by hearing damage. This thesis investigates dynamic, or time-dependent, properties of cochlear gain. A time scale from 0 to 10 ms is considered to ensure that cochlear processing is investigated without including influences from higher stages in the brain. The results are expected
to provide insight into how e.g. onsets of sounds are processed by the intact human system. Click-evoked otoacoustic emissions (CEOAEs) were used as a non-invasive technique to investigate cochlear gain mechanisms. CEOAEs are echoes to click stimuli that can be recorded in the ear canal, and that are produced in the cochlea as a byproduct of the nonlinear gain mechanism. Experimental results demonstrated that the CEOAE level-curve, i.e. the relation between click and CEOAE level, altered when a click was presented close in time to the evoking click. This effect was named "temporal adaptation" of the CEOAE level-curve, and was shown to operate on a time scale of 0 to 8 ms. The relation between dynamic features of CEOAEs and the underlying cochlear gain mechanism was furthermore investigated by means of a numerical model of the cochlea that simulates CEOAEs. The simulations provided insight into level-dependent features of the cochlear gain mechanism that underlie the generation of the CEOAE. In order to account for key features of temporal adaptation in CEOAEs, the existence of a time dependence in the cochlear gain mechanism was suggested. Overall, this study has demonstrated that cochlear compression characteristics can change on a time scale of 0–8 ms. The existence of such a time constant in cochlear compression may be of interest for the future development of signal processing in hearing instruments.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Harvard Medical School
Authors: Verhulst, S. (Intern), Dau, T. (Intern), Harte, J. (Intern), Christopher A., S. (Ekstern)
Number of pages: 180
Publication date: 2010

Publication information
ISBN (Print): 978-87-92465-24-5
Original language: English

Series: CONTRIBUTIONS TO HEARING RESEARCH
Number: 8
Main Research Area: Technical/natural sciences
Electronic versions:
PhDThesisSV[1].pdf
Source: orbit
Source-ID: 275119
Publication: Research › Ph.D. thesis – Annual report year: 2010

Detection and identification of monaural and binaural pitch contours in dyslexic listeners
Binaural pitch stimuli were used in several recent studies to test for the presence of binaural auditory impairment in reading-disabled subjects. The outcome of three of these studies (Dougherty et al., 1998; Edwards et al., 2004; Chait et al., 2007) has been contradictory: Where the former two found that a majority of dyslexic subjects were unable to hear binaural pitch, the latter obtained a clear response of dyslexic listeners to Huggins' pitch (HP) (Cramer and Huggins, 1958). The present study clarified whether impaired binaural pitch perception is found in dyslexia. Results from a pitch contour identification test, performed in 31 dyslexic listeners and 31 matched controls, clearly showed that dyslexics perceived HP as well as the controls. Both groups also showed comparable results with a similar-sounding, monaurally-detectable, pitch-evoking stimulus. However, nine of the dyslexic subjects had difficulty identifying pitch contours, independent of the stimulus used. The ability of subjects to correctly identify pitch contours was found to be significantly correlated to measures of frequency discrimination. This correlation may be attributed to the similarity of the experimental tasks and probably reflects impaired cognitive mechanisms related to auditory memory or auditory attention rather than impaired low-level auditory processing per se.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research, Katholieke Universiteit
Authors: Santurette, S. (Intern), Dau, T. (Intern), Poelmans, H. (Ekstern), Luts, H. (Ekstern), Wouters, J. (Ekstern)
Publication date: 2010

Host publication information
Title of host publication: Proceedings of the International Symposium on Auditory and Audiological Research: Binaural Processing and Spatial Hearing
Main Research Area: Technical/natural sciences
Links:
http://www.isaar.eu
Source: orbit
Source-ID: 263965
Detection and identification of monaural and binaural pitch contours in dyslexic listeners

The use of binaural pitch stimuli to test for the presence of binaural auditory impairment in reading-disabled subjects has so far led to contradictory outcomes. While some studies found that a majority of dyslexic subjects was unable to perceive binaural pitch, others obtained a clear response of dyslexic listeners to Huggins' pitch (HP). The present study clarified whether impaired binaural pitch perception is found in dyslexia. Results from a pitch contour identification test, performed in 31 dyslexic listeners and 31 matched controls, clearly showed that dyslexics perceived HP as well as the controls. Both groups also showed comparable results with a similar-sounding, but monaurally detectable, pitch-evoking stimulus. However, nine of the dyslexic subjects were found to have difficulty identifying pitch contours both in the binaural and the monaural conditions. The ability of subjects to correctly identify pitch contours was found to be significantly correlated to measures of frequency discrimination. This correlation may be attributed to the similarity of the experimental tasks and probably reflects impaired cognitive mechanisms related to auditory memory or auditory attention rather than impaired low-level auditory processing per se.

General information

State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research, Katholieke Universiteit
Authors: Santurette, S. (Intern), Poelmans, H. (Ekstern), Luts, H. (Ekstern), Ghesquière, P. (Ekstern), Wouters, J. (Ekstern), Dau, T. (Intern)
Pages: 515-524
Publication date: 2010
Main Research Area: Technical/natural sciences

Publication information

Journal: J A R O
Volume: 11
Issue number: 3
ISSN (Print): 1525-3961
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 2.61 SJR 1.363 SNIP 1.269
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 1.688 SNIP 1.48 CiteScore 2.95
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 1.592 SNIP 1.371 CiteScore 2.84
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 1.333 SNIP 1.432 CiteScore 2.67
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 1.535 SNIP 1.299 CiteScore 2.74
ISI indexed (2012): ISI indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 1.416 SNIP 1.5 CiteScore 2.98
ISI indexed (2011): ISI indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 1.617 SNIP 1.406
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 1.406 SNIP 1.17
Detection of dynamically varying interaural time differences

Humans are highly sensitive to Interaural Time Differences (ITDs) in stimuli presented via headphones. For broadband noise stimuli of long durations, ITD detection thresholds can be as low as 10 to 15 ms. When the stimulus duration is shortened, thresholds increase by about a factor 2 for a tenfold decrease in duration. ITD thresholds also increase, when the probe carrying an ITD is surrounded by diotic fringes. When a 5-ms probe is combined with preceding or trailing fringes, the effect of a fringe preceding the probe is stronger than that of a trailing fringe for fringe durations <35 ms. The effect of fringes surrounding the probe is equal to the addition of the effects of the individual fringes. In this contribution, we present behavioral data for the same experimental condition, called dynamically varying ITD detection, but for a wider range of probe and fringe durations. Probe durations varied between 5 and 400 ms, and fringe durations had values of 5, 20, 100 or 200 ms. In contrast to earlier findings, we observed for most duration combinations a stronger effect of the trailing fringe than of the preceding fringe. For these configurations, the effect of surrounding fringes was dominated by the trailing fringe. Only for the combination of 5-ms fringes with 5-ms probes did we see the clear dominance of the preceding fringe. These results are not easy to align with the concept of onset emphasis often used to explain binaural localization data for short stimuli. In fact the data seem to be difficult to predict with a purely signal-driven model of perception and thus form an interesting challenge for modeling human localization.

Equivalent threshold sound pressure levels (ETSPL) for Interacoustics DD 45 supra-aural audiometric earphones

This paper describes the determination and results of pure tone Equivalent Threshold Sound Pressure Levels for the Interacoustics DD45 audiometric earphone equipped with standard Model 51 cushions. The size and shape of the DD45 transducer resembles the classical Telephonics TDH 39 earphone. Pure tone hearing threshold measurements were
performed for both ears of 29 test subjects. All audiometric frequencies from 125 Hz to 8 kHz were used. The data are intended for inclusion in future standardized Reference Equivalent Threshold Sound Pressure Levels. The results show that the DD45 may be a good substitute for THD 39.

**General information**
- **State**: Published
- **Organisations**: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
- **Authors**: Poulsen, T. (Intern)
- **Pages**: 850-855
- **Publication date**: 2010
- **Main Research Area**: Technical/natural sciences

**Publication information**
- **Journal**: International Journal of Audiology
- **Volume**: 49
- **Issue number**: 11
- **ISSN (Print)**: 1499-2027
- **Ratings**:
  - BFI (2018): BFI-level 2
  - Web of Science (2018): Indexed yes
  - BFI (2017): BFI-level 2
  - Web of Science (2017): Indexed Yes
  - BFI (2016): BFI-level 2
  - Scopus rating (2016): CiteScore 2.07 SJR 1.289 SNIP 1.245
  - Web of Science (2016): Indexed yes
  - BFI (2015): BFI-level 2
  - Scopus rating (2015): SJR 1.191 SNIP 1.217 CiteScore 1.79
  - BFI (2014): BFI-level 2
  - Scopus rating (2014): SJR 1.3 SNIP 1.273 CiteScore 1.89
  - Web of Science (2014): Indexed yes
  - BFI (2013): BFI-level 2
  - Scopus rating (2013): SJR 1.191 SNIP 1.499 CiteScore 1.94
  - ISI indexed (2013): ISI indexed yes
  - Web of Science (2013): Indexed yes
  - BFI (2012): BFI-level 2
  - Scopus rating (2012): SJR 1.232 SNIP 1.296 CiteScore 1.79
  - ISI indexed (2012): ISI indexed yes
  - Web of Science (2012): Indexed yes
  - BFI (2011): BFI-level 2
  - Scopus rating (2011): SJR 1.29 SNIP 1.209 CiteScore 1.78
  - ISI indexed (2011): ISI indexed yes
  - BFI (2010): BFI-level 2
  - Scopus rating (2010): SJR 1.075 SNIP 1.118
  - Web of Science (2010): Indexed yes
  - BFI (2009): BFI-level 2
  - Scopus rating (2009): SJR 1.179 SNIP 1.11
  - Web of Science (2009): Indexed yes
  - BFI (2008): BFI-level 1
  - Scopus rating (2008): SJR 1.217 SNIP 1.372
  - Web of Science (2008): Indexed yes
  - Scopus rating (2007): SJR 0.999 SNIP 1.071
  - Scopus rating (2006): SJR 0.957 SNIP 1.22
  - Web of Science (2006): Indexed yes
  - Scopus rating (2005): SJR 1.013 SNIP 1.168
  - Scopus rating (2004): SJR 0.594 SNIP 1.132
  - Scopus rating (2003): SJR 0.29 SNIP 0.301
  - Web of Science (2003): Indexed yes
Estimating individual listeners' auditory-filter bandwidth in simultaneous and non-simultaneous masking

Frequency selectivity in the human auditory system is often measured using simultaneous masking of tones presented in notched noise. Based on such masking data, the equivalent rectangular bandwidth (ERB) of the auditory filters can be derived by applying the power spectrum model of masking and assuming a rounded-exponential filter shape. If a forward masking paradigm is used instead of simultaneous masking, filter estimates typically show significantly sharper tuning. This difference in frequency selectivity has commonly been related to spectral suppression mechanisms observed in the cochlea. Considering bandwidth estimates from previous studies based on forward masking, only average data across a number of subjects have been considered. The present study is concerned with bandwidth estimates in simultaneous and forward masking in individual normal-hearing subjects. In order to investigate the reliability of the individual estimates, a statistical resampling method is applied. It is demonstrated that a rather large set of experimental data is required to reliably estimate auditory filter bandwidth, particularly in the case of simultaneous masking. The poor overall reliability of the filter estimates was found to be mainly related to the very short tone duration (i.e., 10 ms) that was chosen. Applying 300-ms long tones in simultaneous masking drastically improved the reliability of the filter estimates. The tone duration in forward masking had to be very short to elicit a sufficient amount of masking. Based on extensive data for three subjects, the difference between forward and simultaneous masking estimates of auditory filter bandwidth was found to be even larger than previously reported, with a bandwidth decrease by a factor of about 1.8 rather than 1.4. The results of the study can be used to optimize the measures of frequency selectivity which is particularly useful when studying consequences of (individual) hearing impairment.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Buchholz, J. (Intern), Caminade, S. (Ekstern), Strelcyk, O. (Intern), Dau, T. (Intern)
Pages: 3213-3219
Publication date: 2010

Host publication information
Title of host publication: Proceedings of the 2010 Annual Conference of the Australian Acoustical Society
Publisher: International Commission for Acoustics
Main Research Area: Technical/natural sciences
Psychoacoustics, Masking, Auditory filters
Source: orbit
Source-ID: 264023
Publication: Research › Article in proceedings – Annual report year: 2010

Externalization versus Internalization of Sound in Normal-hearing and Hearing-impaired Listeners

The externalization of sound, i.e. the perception of auditory events as being located outside of the head, is a natural phenomenon for normal-hearing listeners, when perceiving sound coming from a distant physical sound source. It is potentially useful for hearing in background noise, but the relevant cues might be distorted by a hearing impairment and also by the processing of the incoming sound through hearing aids. In this project, two intuitive tests in natural real-life surroundings were developed, which capture the limits of the perception of externalization. For this purpose, an auralization system for headphones using individual cues was implemented and a strategy to modify the degree of the externalization was proposed. While normal-hearing listeners obtained consistent results, both individually and across subjects, the limits of externalization varied more within and across listeners in the hearing-impaired group. Partly, there was an influence by the direction of sound incidence. On average across subjects, the dynamic range available to perceive externalization was reduced compared to normal-hearing subjects. Overall, it was shown that hearing-impaired listeners are able to perceive externalization, but also that they are less sensitive to minor deviations from complete internalization and externalization.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research, Oticon A/S
Authors: Ohl, B. (Ekstern), Laugesen, S. (Ekstern), Buchholz, J. (Intern), Dau, T. (Intern)
Number of pages: 136
Free-field correction values for Interacoustics DD 45 supra-aural audiometric earphones

This paper report free-field correction values for the Interacoustics DD 45 audiometric earphones. The free-field correction values for earphones provide the loudness based equivalence to loudspeaker presentation. Correction values are especially used for the calibration of audiometric equipment for speech audiometry performed with headphones. Calibration values may be found in e.g. the ISO 389 series of standards. The free-field correction values were determined by means of loudness balance measurements of one-third octave noises (centre frequencies 125 Hz to 8000 Hz) presented alternately from a loudspeaker in a free field and from the earphones. The procedure was essentially in accordance with the free-field frequency response procedure described in IEC 60268-7: Headphones and earphones. The study sample consisted of four earphones and 14 test subjects. Free field correction values are reported for the acoustic coupler IEC 60318-3 (NBS 9-A) and for the ear simulator IEC 60318-1. The results are in good agreement with the results of another independent investigation. The reported free-field correction values may be used as part of the basis for future standardization of the DD 45 earphone.
Impact of regularization of near field coding filters for 2D and 3D higher-order Ambisonics on auditory distance cues

A known challenge in sound field reproduction techniques such as high-order Ambisonics (HOA) is the reproduction of nearby sound sources. In order to reproduce such nearby sound sources, the near-field compensated (NFC) method with angular weighting windows (AWWs) has been previously proposed for HOA [1]. Considering auditory distance perception, (low-frequency) interaural level differences represent the main auditory cue for nearby real sound sources outside the median plane. Simulations showed that these ILD cues can be reproduced with existing weighted NFC-HOA methods for frequencies above about 500 Hz. However, since low-frequency ILD cues seem to be important for distance perception, a novel regularization function is proposed as AWW that can reproduce natural ILDs down to about 250 Hz, even in realistic playback environments. Using this regularization function, a listening test showed an improved distance perception performance for lateral sources compared to frontal ones. This improvement was greater for 3D reproduction than for 2D but slightly lower than for real sources. The distance of virtual nearby sources reproduced by NFC-HOA with the regularization function as AWW can be perceived nearly as accurately as corresponding physical ones.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Favrot, S. E. (Intern), Buchholz, J. (Intern)
Pages: 4
Publication date: 2010

Host publication information
Title of host publication: Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics
Main Research Area: Technical/natural sciences
Conference: International Symposium on Ambisonics and Spherical Acoustics, Paris, France, 01/01/2010
sound field reproduction, distance perception, Higher-order Ambisonics, near field compensated
Electronic versions:
SFJB_AmbSymp10-final.pdf
Links:
http://ambisonics10.ircam.fr/drupal/?q=proceedings
Source: orbit
Source-ID: 262228
Improving hearing aids through listening tests in a virtual sound environment

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Oticon A/S
Authors: Minnaar, P. (Ekstern), Favrot, S. E. (Intern), Buchholz, J. (Intern)
Pages: 40-44
Publication date: 2010
Main Research Area: Technical/natural sciences

Publication information
Journal: Hearing Journal
Volume: 63
Issue number: 10
ISSN (Print): 0745-7472
Ratings:
Scopus rating (2016): SJR 0.123 SNIP 0.012 CiteScore 0.04
Scopus rating (2015): SJR 0.105 SNIP 0 CiteScore 0.02
Scopus rating (2014): SJR 0.127 SNIP 0.164 CiteScore 0.06
Scopus rating (2013): SJR 0.174 SNIP 0.458 CiteScore 0.1
ISI indexed (2013): ISI indexed no
Scopus rating (2012): SJR 0.152 SNIP 0.403 CiteScore 0.12
ISI indexed (2012): ISI indexed no
Scopus rating (2011): SJR 0.214 SNIP 0.193 CiteScore 0.12
ISI indexed (2011): ISI indexed no
Scopus rating (2010): SJR 0.209 SNIP 0.375
Scopus rating (2009): SJR 0.228 SNIP 0.376
Scopus rating (2008): SJR 0.246 SNIP 0.407
Scopus rating (2007): SJR 0.291 SNIP 0.469
Scopus rating (2006): SJR 0.27 SNIP 0.432
Scopus rating (2005): SJR 0.317 SNIP 0.43
Scopus rating (2004): SJR 0.197 SNIP 0.324
Scopus rating (2003): SJR 0.166 SNIP 0.286
Scopus rating (2002): SJR 0.126 SNIP 0.007
Original language: English
Electronic versions:
HearingJournalVSE[1].pdf
DOIs:
10.1097/01.HJ.0000389926.64797.3e
Links:
http://journals.lww.com/thehearingjournal/Fulltext/2010/10000/Improving_hearing_aids_through_listening_tests_in.7.aspx#
Source: orbit
Source-ID: 267925
Publication: Research - peer-review › Journal article – Annual report year: 2010

Investigation of symphony orchestra musicians' use of hearing protectors
A questionnaire study was performed about the use of hearing protectors in Danish symphony orchestras. The musicians in three Danish symphony orchestras were asked to complete a questionnaire about their use of hearing protection. A total of 146 musicians filled in the questionnaire. Results showed that musicians are aware of the dangers of loud music, that they only use hearing protectors to some extent, and that hearing protector is sometimes used only in one ear. Musicians are concerned about their hearing and musicians that experience hearing problems use hearing protectors more frequently. At present an investigation is performed about the use of hearing aids as hearing protectors by symphony orchestra musicians. Preliminary results from this investigation will be presented at the conference.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Aalto University
Authors: Poulsen, T. (Intern), Koskinen, H. (Ekstern)
LoRA: A Loudspeaker-Based Room Auralization System

In order to study basic human perception in reverberant environments, a novel loudspeaker-based room auralization (LoRA) system is proposed in this paper. The LoRA system efficiently combines modern room acoustic models with high-order Ambisonic auralization. An objective evaluation has been carried out demonstrating the applicability of the LoRA system. Room acoustic parameters (reverberation time, clarity, speech transmission index and inter-aural cross correlation coefficients) of room impulse responses were compared at the input and the simulated output of the LoRA system. Results show that the involved signal processing preserves the temporal, spectral and spatial properties of the room impulse response captured by these parameters. This flexible research platform will be useful for studying auditory processing and perception in normal-hearing and hearing-impaired listeners in fully controlled and realistic environments.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering
Authors: Favrot, S. E. (Intern), Buchholz, J. (Intern)
Pages: 364-375
Publication date: 2010
Main Research Area: Technical/natural sciences

Publication information
Journal: Acustica United with Acta Acustica
Volume: 96
Issue number: 2
ISSN (Print): 1610-1928
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.12 SJR 0.451 SNIP 0.834
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.617 SNIP 1.093 CiteScore 1.11
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.615 SNIP 1.071 CiteScore 0.89
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.597 SNIP 1.6 CiteScore 1.05
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.602 SNIP 0.963 CiteScore 0.81
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.515 SNIP 0.918 CiteScore 0.65
Loudspeaker-based room auralization in auditory perception research

Recently a loudspeaker-based room auralisation (LoRA) system has been developed at CAHR, which combines modern room acoustic modeling techniques with high-order Ambisonics auralisation. The environment provides: (i) a flexible research tool to study the signal processing of the normal, impaired, and aided-impaired auditory system in realistic environments and (ii) a framework to evaluate the effect of different room modeling and auralisation methods on auditory perception. The applicability of such environment is demonstrated using different objective room acoustic measures. Different experimental results are presented, including measures of distance perception and the effect of early reflections on speech intelligibility.

General information
State: Published
Organisations: Centre for Applied Hearing Research, Department of Electrical Engineering, Hearing Systems
Authors: Buchholz, J. (Intern), Favrot, S. E. (Intern)
Publication date: 2010

Host publication information
Title of host publication: International Workshop on the principles and applications of spatial hearing
Publisher: IWPASH
Main Research Area: Technical/natural sciences
Conference: International Workshop on the principles and applications of spatial hearing, Sendai, Japan, 01/01/2009
Virtual Auditory Environments, Room Auralization, Hearing Research
Source: orbit
Source-ID: 264028
Publication: Research - Article in proceedings – Annual report year: 2010
Modeling human auditory evoked brainstem responses based on nonlinear cochlear processing

The aim of this study was to accurately simulate auditory evoked potentials (AEPs) from various classical stimuli such as clicks and tones, often used in research and clinical diagnostics. In an approach similar to Dau (2003), a model was developed for the generation of auditory brainstem responses (ABR) to transient sounds and frequency following responses (FFR) to tones. The model includes important cochlear processing stages (Zilany and Bruce, 2006) such as basilar-membrane (BM) tuning and compression, inner hair-cell (IHC) transduction, and IHC auditory-nerve (AN) synapse adaptation. To generate AEPs recorded at remote locations, a convolution was made on an empirically obtained elementary unit waveform with the instantaneous discharge rate function for the corresponding AN unit. AEPs to click-trains, as well as to tone pulses at various frequencies, were both modelled and recorded at different stimulation levels and repetition rates. The observed nonlinearities in the recorded potential patterns, with respect to ABR wave V latencies and amplitudes, could be largely accounted for by level-dependent BM processing as well as effects of short-term neural adaptation. The present study provides further evidence for the importance of cochlear tuning and AN adaptation on AEP patterns, and provides a useful basis for the study of more complex stimuli including speech.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Harte, J. (Intern), Rønne, F. M. (Intern), Dau, T. (Intern)
Publication date: 2010

Host publication information
Title of host publication: Proceedings of the 20th International Congress on Acoustics
Main Research Area: Technical/natural sciences
Electronic versions: p422.pdf
Source: orbit
Source-ID: 263952
Publication: Research - peer-review › Article in proceedings – Annual report year: 2010

Modelling a damaged cochlea: beyond non-speech psychophysics

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Boston University
Authors: Jepsen, M. L. (Intern), Dau, T. (Intern), Ghitza, O. (Ekstern)
Publication date: 2010

Host publication information
Title of host publication: Binaural Processing and Spatial Hearing
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 264013
Publication: Research - peer-review › Article in proceedings – Annual report year: 2010

Monaural and binaural benefit from early reflections for speech intelligibility

The auditory system takes advantage of early reflections (ER’s) in a room by integrating them with the direct sound (DS) and thereby increasing the effective speech level. The energy, spectral content and direction of the ER’s are dependent on wall absorptions and room geometries and therefore different from the DS. By using an ER pattern from a real room and reproducing it in a loudspeaker-based virtual auditory environment, the important characteristics of the ER’s can be preserved and the benefit from ER’s for speech intelligibility (SI) can be quantified. In the present study SI was measured in such a realistic sound field by varying the level of the ER’s and the DS of the speech signal independently. The efficiency of the ER’s was then defined as the ratio of the useful ER energy and the total ER energy at the speech reception threshold (SRT). Furthermore monaural SI measures were performed to investigate if ER processing is a monaural or binaural effect.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering
Authors: Arweiler, I. (Intern), Buchholz, J. (Intern), Dau, T. (Intern)
Publication date: 2010

Host publication information
Predicting effects of impaired cochlear processing on consonant discrimination in stationary noise

Cochlear hearing loss is typically associated with reduced sensitivity due to inner hair-cell (IHC) and outer hair-cell (OHC) dysfunction. OHC dysfunction also leads to supra-threshold deficits, such as reduced basilar-membrane (BM) compression as well as reduced frequency selectivity and temporal resolution. Listeners with a cochlear damage typically have difficulties with speech understanding in the presence of background noise. In this study, the goal was to investigate the relation between individual consonant confusions in stationary noise and deficits in cochlear signal-processing as characterized by the audiogram and estimates of the BM input-output characteristics. Cochlear processing in individual listeners was simulated using a computational model of auditory signal processing and perception (CASP) and was used as a front end in a consonant discrimination system. Individual error patterns from a Diagnostic Rhyme Test (DRT) were measured and analyzed in terms of acoustic-phonetic features. This was done for three listeners with cochlear hearing loss and at two signal-to-noise ratios. It is shown that the predicted errors patterns matched the measured patterns in most conditions. Thus, an incomplete representation of the speech sounds due to deficits in cochlear processing could be related to the performance in the speech perception task. In addition, it was studied to what extent the data could be accounted for based on reduced sensitivity only – assuming that BM compression, frequency selectivity and temporal resolution are the same as in normal-hearing listeners. For two out of the three listeners, the supra-threshold deficits needed to be included in order to account for the data, while for the third listener, the predicted error rates were similar for the two model versions. Overall, the results suggest a clear relation between deficits in cochlear signal processing and consonant identification error patterns, and indicate that the supra-threshold deficits associated with a cochlear damage need to be taken into account. The findings might be interesting for applications, such as the evaluation of hearing-instrument signal processing, where the effects of specific processing strategies can be simulated for individual hearing losses.

Prediction of speech intelligibility based on an auditory preprocessing model

Classical speech intelligibility models, such as the speech transmission index (STI) and the speech intelligibility index (SII) are based on calculations on the physical acoustic signals. The present study predicts speech intelligibility by combining a psychoacoustically validated model of auditory preprocessing [Dau et al., 1997. J. Acoust. Soc. Am. 102, 2892-2905] with a simple central stage that describes the similarity of the test signal with the corresponding reference signal at a level of the internal representation of the signals. The model was compared with previous approaches, whereby a speech in noise experiment was used for training and an ideal binary mask experiment was used for evaluation. All three models were able to capture the trends in the speech in noise training data well, but the proposed model provides a better prediction of the binary mask test data, particularly when the binary masks degenerate to a noise vocoder.
Reproduction of nearby sound sources using high-order Ambisonics: Implementation and evaluation

General information
State: Published
Revisiting perceptual compensation for effects of reverberation in speech identification

Listeners were given the task to identify the stop-consonant [t] in the test-word "stir" when the word was embedded in a carrier sentence. Reverberation was added to the test-word, but not to the carrier, and the ability to identify the [t] decreased because the amplitude modulations associated with the [t] were smeared. When a similar amount of reverberation was also added to the carrier sentence, the listeners' ability to identify the stop-consonant was restored. This phenomenon has in previous research been considered as evidence for an extrinsic compensation mechanism for reverberation in the human auditory system [Watkins (2005). J. Acoust. Soc. Am. 118, 249-262]. In the present study, the reverberant test-word was embedded in additional non-reverberant carriers, such as white noise, speech-shaped noise and amplitude modulated noise. In addition, a reference condition was included where the test-word was presented in isolation, i.e., without any carrier stimulus. In all of these conditions, the ability to identify the stop-consonant [t] was enhanced relative to the condition using the non-reverberant speech carrier. The results suggest that the non-reverberant speech carrier produces an interference effect that impedes the identification of the stop-consonant. These findings raise doubts about the existence of the compensation mechanism.
<table>
<thead>
<tr>
<th>Year</th>
<th>Scopus Rating (SJR)</th>
<th>Scopus Rating (SNIP)</th>
<th>Web of Science</th>
<th>BFI Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>0.763</td>
<td>1.622</td>
<td>yes</td>
<td>2</td>
</tr>
<tr>
<td>2011</td>
<td>0.695</td>
<td>1.642</td>
<td>yes</td>
<td>2</td>
</tr>
<tr>
<td>2010</td>
<td>0.754</td>
<td>1.528</td>
<td>yes</td>
<td>2</td>
</tr>
<tr>
<td>2009</td>
<td>0.783</td>
<td>1.717</td>
<td>yes</td>
<td>2</td>
</tr>
<tr>
<td>2008</td>
<td>0.848</td>
<td>1.633</td>
<td>yes</td>
<td>2</td>
</tr>
<tr>
<td>2007</td>
<td>0.865</td>
<td>1.647</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>0.752</td>
<td>1.559</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>0.954</td>
<td>1.749</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>0.77</td>
<td>1.787</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>0.882</td>
<td>1.712</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>0.87</td>
<td>1.501</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>0.719</td>
<td>1.467</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.621</td>
<td>1.411</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>0.591</td>
<td>1.319</td>
<td>yes</td>
<td></td>
</tr>
</tbody>
</table>

**Original language**: English

**Reverberation, White noise, Amplitude modulation, Speech**

**Electronic versions**:

9A637097d01.pdf

**DOIs**:

10.1121/1.3494508

**Bibliographical note**

Copyright (2010) Acoustical Society of America. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the Acoustical Society of America.

Source: orbit

Source-ID: 264050

Publication: Research - peer-review › Journal article – Annual report year: 2010

**Speech intelligibility enhancement by early reflections for normal-hearing and hearing-impaired listeners**

**General information**

State: Published

Organisations: Hearing Systems, Department of Electrical Engineering

Authors: Arweiler, I. (Intern), Buchholz, J. (Intern), Dau, T. (Intern)

Publication date: 2010

Event: Abstract from 30th International Conference of Audiology, São Paulo, Brazil.

Main Research Area: Technical/natural sciences

Source: orbit
Suitable reverberation time for halls for rock and pop music

The existing body of literature regarding the acoustic design of concert halls has focused almost exclusively on classical music, although there are many more performances of popular music, including rock and pop. Objective measurements were made of the acoustics of 20 rock music venues in Denmark and a questionnaire was used in a subjective assessment of those venues with professional rock musicians and sound engineers as expert listeners. Correlations between the measurements show that clarity, including bass frequencies down to 63 Hz, is important for the general impression of the acoustics of the hall. The best-rated halls in the study have reverberation times that are approximately frequency independent from 0.6 to 1.2 s for hall volumes from 1000 to 6000 m3. The worst rated halls in the study had significantly higher reverberation times in the 63 and 125 Hz bands. Since most audiences at rock concerts are standing, absorption coefficients were measured with a standing audience from 63 Hz to 4 kHz. These measurements showed that a standing audience absorbs about five times as much energy in mid-/high-frequency bands as in low-frequency bands.
Temporal adaptation of the click-evoked otoacoustic emission level-curve reveals dynamic properties of human cochlear processing

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research, Harvard Medical School
Authors: Verhulst, S. (Intern), Harte, J. (Intern), Shera, C. A. (Ekstern), Dau, T. (Intern)
Number of pages: 46
Publication date: 2010

Host publication information
Title of host publication: Abstracts of the thirty-third annual midwinter research meeting
Main Research Area: Technical/natural sciences
Conference: Midwinter Research Meeting of the Association of Research in Otolaryngology, Anaheim, CA, 01/01/2010
Links:
http://www.aro.org

Bibliographical note
Copyright (2010) Acoustical Society of America. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the Acoustical Society of America.

Source: orbit
Source-ID: 256392
Publication: Research - peer-review › Journal article – Annual report year: 2010

Temporal adaptation of the click-evoked otoacoustic emission level-curve reveals dynamic properties of human cochlear processing

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research, Harvard Medical School
Authors: Verhulst, S. (Intern), Harte, J. (Intern), Shera, C. A. (Ekstern), Dau, T. (Intern)
Number of pages: 46
Publication date: 2010

Host publication information
Title of host publication: Abstracts of the thirty-third annual midwinter research meeting
Main Research Area: Technical/natural sciences
Conference: Midwinter Research Meeting of the Association of Research in Otolaryngology, Anaheim, CA, 01/01/2010
Links:
http://www.aro.org

Bibliographical note
Copyright (2010) Acoustical Society of America. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the Acoustical Society of America.

Source: orbit
Source-ID: 256392
Publication: Research - peer-review › Journal article – Annual report year: 2010
The Danish hearing in noise test

Objective: A Danish version of the hearing in noise test (HINT) has been developed and evaluated in normal-hearing (NH) and hearing-impaired (HI) listeners. The speech material originated from Nielsen & Dau (2009) where a sentence-based intelligibility equalization method was presented. Design: In the present study, the speech material was evaluated for naturalness and a subset of sentences selected. The new sentence lists were validated, and after three weeks retested. An additional experiment investigated how recollection of sentences affected the listeners’ performance. Study sample: 16 NH and 16 HI listeners participated in the validation and retest. Twelve HI listeners participated in the experiment on recollection. Results: The average speech recognition threshold in noise (SRT N) for the NH listeners was 2.52 dB, with an overall standard deviation of 0.87 dB. The within-subject standard deviation was similar for the NH and the HI listeners. In the retest, the SRT N decreased by 0.4 dB in both groups. Conclusions: The Danish HINT consists of 10 test lists and three practice lists each containing 20 sentences. The validation results are comparable to those of other versions of HINT. The test seems equally reliable for NH and HI listeners. After three weeks, reliable results can be obtained when sentence lists are reused with the same listeners.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Nielsen, J. B. (Intern), Dau, T. (Intern)
Pages: 202-208
Publication date: 2010
Main Research Area: Technical/natural sciences

Publication Information
Journal: International Journal of Audiology
Volume: 50
Issue number: 3
ISSN (Print): 1499-2027
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 2.07 SJR 1.289 SNIP 1.245
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 1.191 SNIP 1.217 CiteScore 1.79
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 1.3 SNIP 1.273 CiteScore 1.89
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 1.191 SNIP 1.499 CiteScore 1.94
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 1.232 SNIP 1.296 CiteScore 1.79
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 1.29 SNIP 1.209 CiteScore 1.78
ISI indexed (2011): ISI indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 1.075 SNIP 1.118
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 1.179 SNIP 1.11
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 1
Scopus rating (2008): SJR 1.217 SNIP 1.372
The Effect of a Voice Activity Detector on the Speech Enhancement

A multimicrophone speech enhancement algorithm for binaural hearing aids that preserves interaural time delays was proposed recently. The algorithm is based on multichannel Wiener filtering and relies on a voice activity detector (VAD) for estimation of second-order statistics. Here, the effect of a VAD on the speech enhancement of this algorithm was evaluated using an envelope-based VAD, and the performance was compared to that achieved using an ideal error-free VAD. The performance was considered for stationary directional noise and nonstationary diffuse noise interferers at input SNRs from −10 to +5 dB. Intelligibility-weighted SNR improvements of about 20 dB and 6 dB were found for the directional and diffuse noise, respectively. No large degradations (}
The influence of spectral and spatial characteristics of early reflections on speech intelligibility

The auditory system employs different strategies to facilitate speech intelligibility in complex listening conditions. One of them is the integration of early reflections (ER's) with the direct sound (DS) to increase the effective speech level. So far the underlying mechanisms of ER processing have mostly been addressed by manipulating the temporal characteristics of ERs. To achieve a more complete picture, the present study investigates the influence of the spectral and spatial characteristics of ERs on speech intelligibility. Speech intelligibility tests were performed with 9 normal-hearing and 8 hearing-impaired listeners in a virtual auditory environment. In this setup with 29 loudspeakers, the amplitude of the DS and the ERs could be varied independently. The ER pattern was taken from a classroom simulated with the room acoustic software Odeon. Thus, the spectral, spatial and temporal characteristics of the ERs were preserved. The DS of the speech signal was always presented from the front and the ERs were either presented from the front or spatially distributed. Speech intelligibility was measured monaurally and binaurally for different types of interferers. It was found for both groups of listeners that speech intelligibility improved with added ER energy, but less than with added DS energy. An efficiency factor was introduced to quantify this effect. The difference in speech intelligibility could be mainly ascribed to the differences in the spectrum between the speech signals with and without ERs. As the ERs were changed from spatial to frontal presentation the speech intelligibility increased in the same manner for monaural and binaural listening. This indicates that the integration of ERs and DS depends on the direction of the ER's, but not on the listening mode (monaural vs. binaural). The direction-dependency could be explained by the spectral changes introduced by the pinna, head, and torso. The results will be important with regard to the influence of signal processing strategies in modern hearing aids on speech intelligibility, because they might alter the spectral, spatial and temporal cues important for the benefit from ERs.

The Low Pitch of High-Frequency Complex Tones Relies on Temporal Fine Structure Information

High-frequency complex tones containing only unresolved harmonic components with a frequency spacing \( \Delta f \) usually evoke a low pitch equal to \( \Delta f \). However, for inharmonic components, the low pitch is often found to deviate slightly from \( \Delta f \). Whether this pitch shift relies exclusively on temporal fine structure (TFS) cues has been a matter of debate. It is also controversial up to which frequency TFS information remains available, and to what extent envelope cues become dominant as frequency increases. Using a pitch-matching paradigm, this study investigated whether the pitch of transposed tones with unresolved inharmonic components is determined by (A) the time intervals between the most prominent TFS peaks in their waveform (multimodal distribution of matches around subharmonics of the carrier frequency \( f_c \)), (B) the timing between peaks in their envelope (unimodal distribution of matches around the envelope rate \( f_{env} \)), or whether (C) no salient pitch is evoked (random matches). Six musically-trained normal-hearing subjects matched the fundamental pitch of a broadband pulse train to that of transposed tones with carrier frequencies \( f_c = [3, 4, 5, 6, 7] \) kHz and envelope rates \( f_{env} = [f_c/11.5, f_c/14.5] \). All stimuli were presented at 50 dB SPL in broadband pink-noise (13.5 dB/Hz at 1 kHz), and 40 matches per condition were obtained. For \( f_{env} = f_c/11.5 \), the results favored hypothesis A for all values of \( f_c \), indicating that TFS cues are available and used for pitch extraction, up to at least 7 kHz in most subjects. For \( f_{env} = f_c/14.5 \), hypothesis A was valid for values of \( f_c \) up to 5 kHz, and the distribution of matches showed a higher variance indicating a less salient pitch. In other conditions, hypothesis C was valid, suggesting that envelope cues do not take over as TFS cues become unavailable. These results strongly suggest that the monaural representation of TFS persists at high frequencies and that pitch does not rely on envelope coding as such.
Assessment of speech intelligibility in background noise and reverberation

Reliable methods for assessing speech intelligibility are essential within hearing research, audiology, and related areas. Such methods can be used for obtaining a better understanding of how speech intelligibility is affected by, e.g., various environmental factors or different types of hearing impairment. In this thesis, two sentence-based tests for speech intelligibility in Danish were developed. The first test is the Conversational Language Understanding Evaluation (CLUE), which is based on the principles of the original American-English Hearing in Noise Test (HINT). The second test is a modified version of CLUE where the speech material and the scoring rules have been reconsidered. An extensive validation of the modified test was conducted with both normal-hearing and hearing-impaired listeners. The validation showed that the test produces reliable results for both groups of listeners. An important deviation between the two new tests and the original HINT is a new procedure used for equalizing the intelligibility of the speech material during the development process. This procedure produces more accurately equalized sentences than achieved with the original HINT procedure. This study also investigates a fundamentally different method for assessing speech intelligibility. This method is based on the identification of the stop-consonant [t] in a short test-word. The method was originally developed in order to measure the impact of reverberation on speech intelligibility and, in particular, to measure whether the intelligibility of the test-word depends on the reverberation added to a surrounding speech carrier. It has been shown that the intelligibility of a reverberant test-word increases when the same amount of reverberation is also added to the carrier. In the literature, this observation has been interpreted as evidence of an extrinsic compensation mechanism for reverberation in the human auditory system. However, in the present study, it is shown that the listener's perception of the test-word is not only related to the carrier reverberation but also to other of the carrier's acoustic-phonetic properties. The evidence of the extrinsic compensation mechanism is therefore questionable. Overall, the results from the present study may contribute to the development of future speech intelligibility tests in Danish and other languages. The two developed sentence tests are expected to be useful for assessing speech intelligibility with Danish NH and HI listeners.

Spectro-temporal analysis of complex sounds in the human auditory system

Most sounds encountered in our everyday life carry information in terms of temporal variations of their envelopes. These envelope variations, or amplitude modulations, shape the basic building blocks for speech, music, and other complex sounds. Often a mixture of such sounds occurs in natural acoustic scenes, with each of the sounds having its own
characteristic pattern of amplitude modulations. Complex sounds, such as speech, share the same amplitude modulations across a wide range of frequencies. This "comodulation" is an important characteristic of these sounds since it can enhance their audibility when embedded in similar background interferers, a phenomenon referred to as comodulation masking release (CMR). Knowledge of the auditory processing of amplitude modulations provides therefore crucial information for a better understanding of how the auditory system analyses acoustic scenes. The purpose of the present thesis is to develop a computational auditory processing model that accounts for a large variety of experimental data on CMR, in order to obtain a more thorough understanding of the basic processing principles underlying the processing of across-frequency modulations. The second chapter introduces a processing stage, in which information from different peripheral frequency channels is combined. This so-called across-channel processing is assumed to take place at the output of a modulation filterbank, and is crucial in order to account for CMR conditions where the frequency spacing of comodulated components is relatively large. The third chapter investigates the role of nonlinear inner-ear (cochlear) processing on CMR. A compressive non-linearity is incorporated in the modeling framework suggested in the second chapter. This non-linearity is necessary to account for CMR in conditions which are sensitive to cochlear suppression. The fourth chapter examines the role of cognitive processing in different stimulus paradigms: CMR, binaural masking level differences and modulation detection interference are investigated in contexts of auditory grouping. It is shown that auditory grouping can influence the results in conditions where the processing in the auditory system is dominated by across-channel comparisons. Overall, this thesis provides insights into the specific mechanisms involved in the perception of comodulated sounds. The results are important as a basis for future models of complex modulation processing in the human auditory system.

Peripheral auditory processing and speech reception in impaired hearing
One of the most common complaints of people with impaired hearing concerns their difficulty with understanding speech. Particularly in the presence of background noise, hearing-impaired people often encounter great difficulties with speech communication. In most cases, the problem persists even if reduced audibility has been compensated for by hearing aids. It has been hypothesized that part of the difficulty arises from changes in the perception of sounds that are well above hearing threshold, such as reduced frequency selectivity and deficits in the processing of temporal fine structure (TFS) at the output of the inner-ear (cochlear) filters. The purpose of this work was to investigate these aspects in detail. One chapter studies relations between frequency selectivity, TFS processing, and speech reception in listeners with normal and impaired hearing, using behavioral listening experiments. While a correlation was observed between monaural and binaural TFS-processing deficits in the hearing-impaired listeners, no relation was found between TFS processing and frequency selectivity. TFS processing was correlated with speech reception in background noise. Two following chapters investigate cochlear response time (CRT) as an important aspect of the cochlear response to incoming sounds, using objective and behavioral methods. Alterations in CRT were observed for hearing-impaired listeners. A good correspondence between objective and behavioral estimates of CRT indicated that a behavioral lateralization method may be useful for studying spatiotemporal aspects of the cochlear response in human listeners. Behaviorally estimated filter bandwidths accounted for the observed alterations of CRTs in the hearing-impaired listeners, i.e., CRT was found to be inversely related to individual filter bandwidth. Overall, this work provides insights into factors affecting auditory processing in listeners with impaired hearing and may have implications for future models of impaired auditory signal processing as well as advanced compensation strategies.
A loudspeaker-based room auralization system for auditory perception research

Most research on basic auditory function has been conducted in anechoic or almost anechoic environments. The knowledge derived from these experiments cannot directly be transferred to reverberant environments. In order to investigate the auditory signal processing of reverberant sounds, a loudspeaker-based room auralisation (LoRA) system is proposed here. The LoRA system efficiently combines modern room acoustic modelling techniques with higher-order Ambisonic auralisation. Thereby, aspects of the auditory precedence effect are utilized to realise highly authentic room reverberation. This system provides a flexible research platform for conducting auditory experiments with normal-hearing, hearing-impaired, and aided hearing-impaired listeners in a fully controlled and realistic environment. This includes measures of basic auditory function (e.g., signal detection, distance perception) and measures of speech intelligibility. A battery of objective tests (e.g., reverberation time, clarity, interaural correlation coefficient) and subjective tests (e.g., speech reception thresholds) is presented that demonstrates the applicability of the LoRA system.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Buchholz, J. (Intern), Favrot, S. E. (Intern)
Pages: 1295-1298
Publication date: 2009

Host publication information
Title of host publication: Proceedings of the International Conference on Acoustics
Volume: 2009
Main Research Area: Technical/natural sciences
Virtual acoustics
Links:
http://www.nag-daga.nl
Source: orbit
Source-ID: 250411
Publication: Research › Article in proceedings – Annual report year: 2009

An Efficient Algorithm for the Discrete Gabor Transform using full length Windows

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Søndergaard, P. L. (Intern)
Publication date: 2009
Conference: The international conference on Sampling Theory and Applications, Marseilles, France, 01/01/2009
Main Research Area: Technical/natural sciences

Publication information
Journal: Sampling Theory in Signal and Image Processing
ISSN (Print): 1530-6429
Ratings:
BFI (2018): BFI-level 1
BFI (2017): BFI-level 1
BFI (2016): BFI-level 1
Scopus rating (2016): SJR 0.305 SNIP 0.609 CiteScore 0.6
BFI (2015): BFI-level 1
Comparison of cochlear delay estimates using otoacoustic emissions and auditory brainstem responses

Different attempts have been made to directly measure frequency specific basilar membrane (BM) delays in animals, e.g., laser velocimetry of BM vibrations and auditory nerve fiber recordings. The present study uses otoacoustic emissions (OAEs) and auditory brainstem responses (ABRs) to estimate BM delay non-invasively in normal-hearing humans. Tone bursts at nine frequencies from 0.5 to 8 kHz served as stimuli, with care taken to quantify possible bias due to the use of tone bursts with different rise times. BM delays are estimated from the ABR latency estimates by subtracting the neural and synaptic delays. This allows a comparison between individual OAE and BM delays over a large frequency range in the same subjects, and offers support to the theory that OAEs are reflected from a tonotopic place and carried back to the cochlear base via a reverse traveling wave.
Detection and identification of monaural and binaural pitch contours in dyslexic listeners

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research, Katholieke Universiteit

Bibliographical note
Copyright 2009. Acoustical Society of America. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the Acoustical Society of America.

Source: orbit
Source-ID: 250409
Publication: Research - peer-review › Journal article – Annual report year: 2009

Electronic versions:
Harte.pdf
DOIs:
10.1121/1.3168508
Development of a Danish speech intelligibility test

Abstract A Danish speech intelligibility test for assessing the speech recognition threshold in noise (SRTN) has been developed. The test consists of 180 sentences distributed in 18 phonetically balanced lists. The sentences are based on an open word-set and represent everyday language. The sentences were equalized with respect to intelligibility to ensure uniform SRTN assessments with all lists. In contrast to several previously developed tests such as the hearing in noise test (HINT) where the equalization is based on scored (objective) measures of word intelligibility, the present test used an equalization method based on subjective assessments of the sentences. The new equalization method is shown to create lists with less variance between the SRTNs than the traditional method. The number of sentence levels included in the SRTN calculation was also evaluated and differs from previous tests. The test was verified with 14 normal-hearing listeners; the overall SRTN lies at a signal-to-noise ratio of -3.15 dB with a standard deviation of 1.0 dB. The list-SRTNs deviate less than 0.5 dB from the overall mean.
Digital Signal Processing for Hearing Instruments

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering
Pages: 898576
Publication date: 2009
Main Research Area: Technical/natural sciences

Publication information
Journal: Eurasip Journal on Advances in Signal Processing
ISSN (Print): 1687-6172
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 1
Scopus rating (2016): SJR 0.313 SNIP 0.78 CiteScore 1.21
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 0.279 SNIP 0.592 CiteScore 0.83
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 0.229 SNIP 0.54 CiteScore 0.7
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 0.267 SNIP 0.506 CiteScore 0.63
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): SJR 0.278 SNIP 0.582 CiteScore 0.72
ISI indexed (2012): ISI indexed yes
BFI (2011): BFI-level 1
Scopus rating (2011): SJR 0.371 SNIP 0.724 CiteScore 0.91
ISI indexed (2011): ISI indexed yes
BFI (2010): BFI-level 1
Scopus rating (2010): SJR 0.403 SNIP 0.982
BFI (2009): BFI-level 1
Scopus rating (2009): SJR 0.474 SNIP 0.823
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 0.468 SNIP 0.897
Scopus rating (2007): SJR 0.386 SNIP 0.913
Scopus rating (2006): SJR 0.362 SNIP 0.92
Scopus rating (2005): SJR 0.519 SNIP 0.968
Scopus rating (2004): SJR 0.603 SNIP 1.155
Scopus rating (2003): SJR 0.63 SNIP 1.023
Scopus rating (2002): SJR 0.14 SNIP 0.329
Scopus rating (2001): SJR 0.118 SNIP 0.372
Scopus rating (2000): SJR 0.115 SNIP 0.236
Scopus rating (1999): SJR 0.194 SNIP 0.381
Original language: English
DOIs:
10.1155/2009/898576
Source: orbit
Source-ID: 263704
Publication: Research - peer-review › Editorial – Annual report year: 2009

Distance Perception in a loudspeaker-based room auralization system
A loudspeaker-based room auralization (LoRA) system has been recently proposed that efficiently combines modern room acoustic modeling techniques with high-order Ambisonics (HOA) auralization to generate virtual auditory environments (VAEs). The reproduction of the distance of sound events in such VAE is very important for its fidelity. A direct-scaling distance perception experiment was conducted to evaluate the LoRA system including the use of near-field control (NFC) for HOA. Experimental results showed that (i) loudspeaker-based auralization in the LoRA system provides similar distance perception to that of the corresponding real environment and that (ii) NFC-HOA provides a significant increase in the range of perceived distances for near sound sources as compared to standard HOA.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering
Authors: Favrot, S. E. (Intern), Buchholz, J. (Intern)
Pages: 7854
Publication date: 2009

Host publication information
Title of host publication: Audio Engineering Society Convention Papers
Place of publication: New York, USA
Publisher: Praesens Verlag
ISBN (Print): 9780937803714
Main Research Area: Technical/natural sciences
High-order Ambisonics, Distance perception, loudspeaker-based room auralization, Virtual auditory environment
Electronic versions:
SFJB_AEsPaper127.pdf
Links:
http://www.aes.org/e-lib/browse.cfm?elib=15049
Source: orbit
Source-ID: 251571
Publication: Research › Article in proceedings – Annual report year: 2009
Effects of pulsing of the target tone on the audibility of partials in inharmonic complex tones

The audibility of partials was measured for complex tones with partials uniformly spaced on an ERBN-number scale. On each trial, subjects heard a sinusoidal "probe" followed by a complex tone. The probe was mistuned downwards or upwards (at random) by 3% or 4.5% from the frequency of one randomly selected partial in the complex (the "target"). The subject indicated whether the target was higher or lower in frequency than the probe. The probe and the target were pulsed on and off and the ramp times and inter-pulse intervals were systematically varied. Performance was better for longer ramp times and longer inter-pulse intervals. In a second experiment, the ability to detect which of two complex tones contained a pulsed partial was measured. The pattern of results was similar to that for experiment 1. A model of auditory processing including an adaptation stage was able to account for the general pattern of the results of experiment 2. The results suggest that the improvement in ability to hear out a partial in a complex tone produced by pulsing that partial is partly mediated by a release from adaptation produced by the pulsing, and does not result solely from reduction of perceptual confusion.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Moore, B. C. (Ekstern), Glasberg, B. R. (Ekstern), Jepsen, M. L. (Intern)
Pages: 3194-3204
Publication date: 2009
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of the Acoustical Society of America
Volume: 125
Issue number: 5
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Equivalent threshold sound pressure levels (ETSPL) for Sennheiser HDA 280 supra-aural audiometric earphones in the frequency range 125 Hz to 8000 Hz

Hearing threshold sound pressure levels were measured for the Sennheiser HDA 280 audiometric earphone. Hearing thresholds were measured for 25 normal hearing test subjects at the 11 audiometric test frequencies from 125 Hz to 8000 Hz. Sennheiser HDA 280 is a supra-aural earphone that may be seen as a substitute for the classical Telephonics TDH 39. The results are given as the Equivalent Threshold Sound Pressure Level, ETSPL, measured in an acoustic coupler specified in IEC 60318-3. The results are in good agreement with an independent investigation from PTB, Braunschweig, Germany. From acoustic laboratory measurements, ETSPL values are calculated for the ear simulator specified in IEC 60318-1. Fitting of earphone and coupler is discussed. The data may be used for a future update of the RETSPL standard for supra-aural audiometric earphones ISO 389-1.
Estimation of cochlear response times using lateralization of frequency-mismatched tones

Behavioral and objective estimates of cochlear response times CRTs and traveling-wave TW velocity were compared for three normal-hearing listeners. Differences between frequency-specific CRTs were estimated via lateralization of pulsed tones that were interaurally mismatched in frequency, similar to a paradigm proposed by Zerlin 1969. J. Acoust. Soc. Am. 46, 1011–1015. In addition, derived-band auditory brainstem responses were obtained as a function of derived-band center frequency. The latencies extracted from these responses served as objective estimates of CRTs. Estimates of TW velocity were calculated from the obtained CRTs. The correspondence between behavioral and objective estimates of CRT and TW velocity was examined. For frequencies up to 1.5 kHz, the behavioral method yielded reproducible results, which were consistent with the objective estimates. For higher frequencies, CRT differences could not be estimated with the behavioral method due to limitations of the lateralization paradigm. The method might be useful for studying the
spatiotemporal cochlear response pattern in human listeners.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Centre for Applied Hearing Research
Authors: Strelcyk, O. (Intern), Dau, T. (Intern)
Pages: 1302-1311
Publication date: 2009
Main Research Area: Technical/natural sciences

Publication information
Volume: 126
Issue number: 3
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 0.848 SNIP 1.633
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 0.865 SNIP 1.647
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 0.752 SNIP 1.559
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 0.954 SNIP 1.749
Web of Science (2005): Indexed yes
Fishing for meaningful units in connected speech

In many branches of spoken language analysis including ASR, the set of smallest meaningful units of speech is taken to coincide with the set of phones or phonemes. However, fishing for phones is difficult, error-prone, and computationally expensive. We present an experiment, based on machine learning, with an alternative approach. Instead of stipulating a basic set of target units, the determination of the set is considered to be part of the learning task. Given 18 recordings of Danish talkers performing a simple lab task, our algorithm produced a set of acoustically well-defined units sufficient for identifying all the major semantic elements (be they parts of words, words or several words), relevant to the task. As the sound encoding used was very simple – fundamental frequency (F0), Harmonicity-to-Noise-Ratio (HNR), and Intensity samples only – the computational complexity involved was far lower than for phonemic recognition. Our findings show that it is possible to automatically characterize a linguistic message, without detailed spectral information or presumptions about the target units. Further, fishing for simple meaningful cues and enhancing these selectively would potentially be a more effective way of achieving intelligibility transfer, which is the end goal for speech transducing technologies.
Frequency Selective Filtering of the Modulation Spectrum and its Impact on Consonant Identification

The spectro-temporal coding of Danish consonants was investigated using an information-theoretic approach. Listeners were asked to identify eleven different consonants spoken in a CV[l] syllable context (where C refers to the initial consonant, V refers to one of three vowels, [l, a, u], and [l] refers to the syllable-final liquid segment). Each syllable was processed so that only a portion of the original audio spectrum was present. Narrow (three-quarter octave) bands of speech, with center frequencies of 750 Hz, 1500 Hz and 3000 Hz, were presented individually and in combination with each other. The modulation spectrum of each band was low-pass filtered at 24, 12, 6 and 3 Hz. Confusion matrices of the consonant-identification data were computed, and from these the amount of information transmitted for each of three phonetic feature dimensions – voicing, manner and place of articulation – was calculated for each condition. This form of analysis provides a simple means of determining whether information associated with each phonetic feature dimension combines linearly across the audio spectrum, and, if not, delineates a method for characterizing the (non-linear) nature of information integration. In addition, the analysis provides a means to associate specific portions of the modulation spectrum with phonetic feature properties. Such analyses indicate that: (1) Accurate, robust decoding of place-of-articulation information requires broadband cross-spectral integration (2) Place-of-articulation information is associated most closely with the modulation spectrum above 6 Hz, with the most significant contribution coming from the region above 12 Hz. (3) Place-of-articulation information is crucial for accurate consonant recognition. Hence, consonant decoding requires cross-spectral integration of the modulation spectrum above 8 Hz. (4) Voicing is mainly associated with the modulation spectrum between 3 and 6 Hz (with a smaller contribution made by the region above 12 Hz). (5) Manner of articulation is most closely associated with the portion of the modulation spectrum above 12 Hz. This form of information-theoretic analysis can be used to delineate those parts of the speech signal of greatest importance for encoding phonetic features associated with intelligibility and speech understanding.

Importance of temporal fine structure information for the low pitch of high-frequency complex tones

General information
State: Published
Organisations: Centre for Applied Hearing Research, Department of Electrical Engineering, Hearing Systems
Authors: Santurette, S. (Intern)
Publication date: 2009
Laboratory determination of annoyance of low frequency noise

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Poulsen, T. (Intern)
Pages: 100-116
Publication date: 2009

Host publication information
Title of host publication: Conference on the improvement of acoustical environment in building : Acoustical Materials Association of Korea
Volume: Vol 4, no 1
Publisher: Acoustical Materials Association of Korea
Main Research Area: Technical/natural sciences
Conference: Conference on the improvement of acoustical environment in building, Gwangju, South Korea, 01/01/2009
Source: orbit
Source-ID: 265698
Publication: Research - peer-review › Article in proceedings – Annual report year: 2009

Models for the Dynamics of Articulatory Lip Movements

General information
State: Published
Organisations: Centre for Applied Hearing Research, Department of Electrical Engineering, Hearing Systems
Authors: Bothe, H. (Intern)
Publication date: 2009

Host publication information
Title of host publication: Proc. 2nd Int. Conf. on Information and Communication Technologies and Accessibility
Publisher: Springer
Main Research Area: Technical/natural sciences
Conference: The 2nd International Conference on Information and Communication Technologies and Accessibility, Hammamet, Tunesia, 01/01/2009
Source: orbit
Source-ID: 250670
Publication: Research - peer-review › Article in proceedings – Annual report year: 2009

Nachbildung von Störungen der Hör- und Sehbahn für Studien an technischen Kommunikationshilfen

General information
State: Published
Organisations: Centre for Applied Hearing Research, Department of Electrical Engineering, Hearing Systems
Authors: Al-Hamdani, S. (Ekstern), Bothe, H. (Intern)
Pages: 262-269
Publication date: 2009

Host publication information
Title of host publication: Proc. 20th Int. Conf. on Electronic Processing of Speech Signals
Publisher: T U Dredsen
Main Research Area: Technical/natural sciences
Conference: The 20th Int. Conf. on Electronic Processing of Speech Signals, Dresden, Germany, 01/01/2009
Source: orbit
Source-ID: 250667
Publication: Research - peer-review › Article in proceedings – Annual report year: 2009

Perceptual effects of Ambisonics on Room Auralization
Precise Stimulation in Auditory Neuroimplants

Processing of Binaural Pitch Stimuli in Hearing-Impaired Listeners
Recent concepts and challenges in hearing research

In everyday life, the speech we listen to is often mixed with many other sound sources as well as reverberation. In such a situation, normal-hearing listeners are able to effortlessly segregate a single voice out of the background, which is commonly known as the 'cocktail party effect'. Conversely, hearing-impaired people have great difficulty understanding speech when more than one person is talking, even when reduced audibility has been fully compensated for by a hearing aid. As with the hearing impaired, the performance of automatic speech recognition systems deteriorates dramatically with additional sound sources. The reasons for these difficulties are not well understood. Only by obtaining a clearer understanding of the auditory system’s coding strategies will it be possible to design intelligent compensation algorithms for hearing devices. This presentation highlights recent concepts of the signal processing strategies employed by the normal as well as impaired auditory system. The aim is to develop a computational auditory signal-processing model, capable of describing the transformation from the acoustical input signal into its "internal (neural) representations". Several stages of processing are considered to be important for a robust signal representation and a deficiency in any of these processing stages is likely to result in a deterioration of the entire system’s performance. A state-of-the-art model of auditory signal processing would be of major practical significance for technical applications, in digital hearing aids, cochlear implants, speech and audio coding, and automatic speech recognition.

General information
State: Published
Organisations: Centre for Applied Hearing Research, Department of Electrical Engineering, Hearing Systems
Authors: Dau, T. (Intern)
Publication date: 2009
Main Research Area: Technical/natural sciences

Bibliographical note
plenary talk
Source: orbit
Source-ID: 250533
Publication: Research › Conference abstract for conference – Annual report year: 2009

Relation between derived-band auditory brainstem response latencies and behavioral frequency selectivity

Derived-band click-evoked auditory brainstem responses ABRs were obtained for normal-hearing NH and sensorineurally hearing-impaired HI listeners. The latencies extracted from these responses, as a function of derived-band center frequency and click level, served as objective estimates of cochlear response times. For the same listeners, auditory-filter bandwidths at 2 kHz were estimated using a behavioral notched-noise masking paradigm. Generally, shorter derived-band latencies were observed for the HI than for the NH listeners. Only at low click sensation levels, prolonged latencies were obtained for some of the HI listeners. The behavioral auditory-filter bandwidths accounted for the across-listener variability in the ABR latencies: Cochlear response time decreased with increasing filter bandwidth, consistent with linear-system theory. The results link cochlear response time and frequency selectivity in human listeners and offer a window to better understand how hearing impairment affects the spatiotemporal cochlear response pattern.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Centre for Applied Hearing Research
Authors: Strelcyk, O. (Intern), Christoforidis, D. (Intern), Dau, T. (Intern)
Pages: 1878-1888
Publication date: 2009
Main Research Area: Technical/natural sciences

Publication information
Volume: 126
ISSN (Print): 0001-4966
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing

Frequency selectivity, temporal fine-structure (TFS) processing, and speech reception were assessed for six normal-hearing (NH) listeners, ten sensorineurally hearing-impaired (HI) listeners with similar high-frequency losses, and two listeners with an obscure dysfunction (OD). TFS processing was investigated at low frequencies in regions of normal hearing, through measurements of binaural masked detection, tone lateralization, and monaural frequency modulation (FM) detection. Lateralization and FM detection thresholds were measured in quiet and in background noise. Speech
reception thresholds were obtained for full-spectrum and lowpass-filtered sentences with different interferers. Both the HI listeners and the OD listeners showed poorer performance than the NH listeners in terms of frequency selectivity, TFS processing, and speech reception. While a correlation was observed between the monaural and binaural TFS-processing deficits in the HI listeners, no relation was found between TFS processing and frequency selectivity. The effect of noise on TFS processing was not larger for the HI listeners than for the NH listeners. Finally, TFS-processing performance was correlated with speech reception in a two-talker background and lateralized noise, but not in amplitude-modulated noise. The results provide constraints for future models of impaired auditory signal processing.
Simulation of Auditory-Visual

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Al-Hamdani, S. (Ekstern), Bothe, H. (Intern)
Publication date: 2009

Host publication information
Title of host publication: Proceedings of the International Conference on Assistive Technologies for People with Vision and Hearing Impairments
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 249332
Publication: Research › Conference article – Annual report year: 2009

Speech intelligibility enhancement by early reflections

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Arweiler, I. (Intern), Buchholz, J. (Intern), Dau, T. (Intern)
Publication date: 2009
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 264051
Publication: Research › Poster – Annual report year: 2009
Strategies and Results for the Evaluation of the Naturalness of the LIPPS Facial Animation System

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Eger, J. (Ekstern), Bothe, H. (Intern)
Publication date: 2009

Host publication information
Title of host publication: Proc. 8th International Conference on Audio-Visual Speech Processing
Main Research Area: Technical/natural sciences
Conference: The 8th International Conference on Audio-Visual Speech Processing, Norwich, United Kingdom, 01/01/2009
Source: orbit
Source-ID: 250672
Publication: Research › Article in proceedings – Annual report year: 2009

Temporal Adaptation in Click-Evoked Otoacoustic Emissions

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Verhulst, S. (Intern), Harte, J. M. (Intern), Dau, T. (Intern)
Publication date: 2009
Event: Abstract from Midwinter Research Meeting of the Association for Research in Otolaryngology, Baltimore, ML, .
Main Research Area: Technical/natural sciences
Links:
Source: orbit
Source-ID: 250425
Publication: Research › Conference abstract for conference – Annual report year: 2009

TESTPERSON OPERATED 2-ALTERNATIVE FORCED CHOICE AUDIOMETRY COMPARED TO TRADITIONAL AUDIOMETRY

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, University of Southern Denmark
Authors: Schmidt, J. H. (Ekstern), Brandt, C. (Ekstern), Jacob Christensen-Dalsgaard, J. (Ekstern), Andersen, T. (Ekstern), Bælum, J. (Ekstern), Poulsen, T. (Intern)
Publication date: 2009
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 265584
Publication: Research - peer-review › Poster – Annual report year: 2009

Validation of a loudspeaker-based room auralization system using speech intelligibility measures

A novel loudspeaker-based room auralization (LoRA) system has been proposed to generate versatile and realistic virtual auditory environments (VAEs) for investigating human auditory perception. This system efficiently combines modern room acoustic models with loudspeaker auralization using either single loudspeaker or high-order Ambisonics (HOA) auralization. The LoRA signal processing of the direct sound and the early reflections was investigated by measuring the speech intelligibility enhancement by early reflections in diffuse background noise. Danish sentences were simulated in a classroom and the direct sound and each early reflection were either auralized with a single loudspeaker, HOA or first-order Ambisonics. Results indicated that (i) absolute intelligibility scores are significantly dependent on the reproduced technique and that (ii) early reflections reproduced with HOA provide a similar benefit on intelligibility as when reproduced with a single loudspeaker. It is concluded that speech intelligibility experiments can be carried out with the LoRA system either with the single loudspeaker or HOA technique.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Deriving cochlear delays in humans using otoacoustic emissions and auditory evoked potentials

A great deal of the processing of incoming sounds to the auditory system occurs within the cochlea. The organ of Corti within the cochlea has differing mechanical properties along its length that broadly gives rise to frequency selectivity. Its stiffness is at maximum at the base and decreases towards the apex, resulting in locally resonant behaviour. This means high frequencies have maximal response at the base and low frequencies at the apex. The wave travelling along the basilar membrane has a longer travel time for low-frequency stimulus than for high-frequency stimulus. The intrinsic relation between frequency and travel time in the cochlea defines the cochlear delay. This delay is directly associated with the signal analysis occurring in the inner ear and is therefore of primary interest to get a better knowledge of this organ. It is possible to estimate the cochlear delay by direct and invasive techniques, but these disrupt the normal functioning of the cochlea and are usually conducted in animals. In order to obtain an estimate of the cochlear delay that is closer to the normally functioning human cochlea, the present project investigates non-invasive methods in normal hearing adults. These methods include: otoacoustic emissions (OAEs), auditory brainstem responses (ABRs) and auditory steady-state responses (ASSRs). A comparison between the three methods was made across and within subjects, in order to highlight the impact of inter-subject variability on the cochlear delay estimates. The estimates of the cochlear delay obtained with OAEs, ABRs and ASSRs were in good agreement with previously reported studies. The comparison between OAE and ABR latency estimates was made over a broader range of frequencies (0.5-8 kHz), compared to previous studies. Below about 2 kHz the OAE delay is twice the cochlear delay, as if the travelling wave went back and forth in the cochlea, as predicted in current theories of OAE generation. This relation, however, does not hold for higher frequencies, calling into question the physical relation between OAE and ABR delay estimates. The comparison between ABR and ASSR latency estimates demonstrated similar rates of latency decrease as a function of frequency. It was further concluded, in this thesis, that OAE measurements are the most appropriate to estimate cochlear delays, since they had the best repeatability and the shortest recording time. Preliminary results are also given for an experiment using stimuli designed to compensate for OAE delays. These were designed to try and reproduce the success of similar stimuli now used routinely to improve ABR signal-to-noise ratio.
Amplitude modulation depth discrimination in hearing-impaired and normal-hearing listeners

The processing of amplitude modulations (AM) of sounds is assumed to be crucial for decoding and understanding of speech in humans. Since hearing-impaired (HI) listeners often suffer from severely hampered speech intelligibility, particularly in reverberant or noisy environments, they might also show degraded performance in AM processing tasks. However, several studies indicated a similar or even better performance in AM detection tasks for sensorineural HI listeners than for normal hearing (NH) listeners when reduced audibility was compensated. In addition to AM detection, this study investigates the differential processing of amplitude modulation depth in HI and NH listeners. AM-depth discrimination of a 4-, 8-, and 30-Hz sinusoidal AM, imposed on a 1- or 4-kHz pure-tone carrier, was measured. The AM of the standard ranged from being well detectable to near threshold. AM-depth discrimination thresholds strongly varied among HI listeners and were elevated in comparison to NH for high standard depths. A model of AM processing is suggested incorporating an individually adjusted simulation of the auditory periphery. To account for the data of HI listeners, however, the key element appeared to be an increased internal noise in the AM-depth domain. Consequences for speech perception are discussed.
A virtual auditory environment for investigating the auditory signal processing of realistic sounds

A loudspeaker-based virtual auditory environment (VAE) has been developed to provide a realistic versatile research environment for investigating the auditory signal processing in real environments, i.e., considering multiple sound sources and room reverberation. The VAE allows a full control of the acoustic scenario in order to systematically study the auditory processing of reverberant sounds. It is based on the ODEON software, which is state-of-the-art software for room acoustic simulations developed at Acoustic Technology, DTU. First, a MATLAB interface to the ODEON software has been developed. Afterwards, different multi-channel playback algorithms have been implemented in MATLAB in order to be able to present a succession of sounds through the VAE in a fast and efficient way. A set of both objective (physical) and subjective (perception-based) measures has been selected to validate the environment and assess its quality.

General information

State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Favrot, S. E. (Intern)
Publication date: 2008
Event: Abstract from Audio signal processing network in Denmark : Seminar on Spatial Sound Processing, Oticon Research Center, Eriksholm, Denmark.
Main Research Area: Technical/natural sciences
Spatial audio
Links:

http://www.asip-net.dk
Estimating the basilar-membrane input-output function in normal-hearing and hearing-impaired listeners

To partly characterize the function of cochlear processing in humans, the basilar membrane (BM) input-output function can be estimated. In recent studies, forward masking has been used to estimate BM compression. If an on-frequency masker is processed compressively, while an off-frequency masker is transformed more linearly, the ratio between the slopes of growth of masking (GOM) functions provides an estimate of BM compression at the signal frequency. In this study, this paradigm is extended to also estimate the knee-point of the I/O-function between linear processing at low levels and compressive processing at medium levels. If a signal can be masked by a low-level on-frequency masker such that signal and masker fall in the linear region of the I/O-function, then a steeper GOM function is expected. The knee-point can then be estimated in the input level region where the GOM changes significantly. Data were collected from eight normal-hearing (NH) and five hearing-impaired (HI) listeners with mild to moderate sensorineural hearing loss. Both groups showed large inter-subject but low intrasubject variability. When the knee-point could be estimated for the HI listeners it was shifted towards higher input levels and compression was similar to that of NH listeners.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Jepsen, M. L. (Intern), Dau, T. (Intern)
Publication date: 2008
Event: Abstract from Acoustics'08, Paris, France.
Main Research Area: Technical/natural sciences
Electronic versions: Jepsen.pdf

Bibliographical note
Copyright (2008) Acoustical Society of America. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the Acoustical Society of America.

Fine-structure processing, frequency selectivity and speech perception in hearing-impaired listeners

Hearing-impaired people often experience great difficulty with speech communication when background noise is present, even if reduced audibility has been compensated for. Other impairment factors must be involved. In order to minimize confounding effects, the subjects participating in this study consisted of groups with homogeneous, symmetric audiograms. The perceptual listening experiments assessed the intelligibility of full-spectrum as well as low-pass filtered speech in the presence of stationary and fluctuating interferers, the individual’s frequency selectivity and the integrity of temporal fine-structure processing. The latter was addressed in a binaural and a monaural experiment. In the binaural experiment, the lateralization threshold was measured for low-frequency tones with ongoing interaural phase delays. In the monaural experiment, detection thresholds for low-rate frequency modulation were obtained. In addition, these binaural and monaural thresholds were measured in a stationary background noise in order to assess the persistence of the fine-structure processing to interfering noise. Apart from elevated speech reception thresholds, the hearing impaired listeners showed poorer performance than the normally hearing in terms of frequency selectivity and fine-structure processing, despite normal audiometric thresholds at the test frequencies. However, the binaural fine-structure processing was not found to be particularly vulnerable to interfering noise in these listeners.

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Strelcyk, O. (Intern), Dau, T. (Intern)
Pages: 3712-3712
Publication date: 2008
Main Research Area: Technical/natural sciences

Publication information
Volume: 123
Issue number: 5
ISSN (Print): 0001-4966
Ratings: BFI (2018): BFI-level 2
Influence of the task of the listener on preference for gain at soft input levels

General information
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems, Centre for Applied Hearing Research
Authors: Connor, H. (Intern), Poulsen, T. (Intern)
Pages: 559-567
Publication date: 2008

Host publication information
Title of host publication: Auditory signal processing in hearing-impaired listeners
Publisher: Center Tryk A/S
ISBN (Print): 87-990013-1-4
Main Research Area: Technical/natural sciences
Links:
http://www.isaar.eu
Source: orbit
Source-ID: 259649
Publication: Research › Article in proceedings – Annual report year: 2007

Modelling across–and within–channel mechanisms in comodulation masking release
The audibility of a target sound embedded in another masking sound can be improved by adding sound energy that is remote in frequency from both the masker and the target. This effect is known as comodulation masking release (CMR) [1] and is observed when the remote sound and the masker share coherent patterns of amplitude modulation. While a large body of data has been presented, the mechanisms underlying CMR are not clear. Neuronal suppression at a cochlear level, the detection of modulation beatings within auditory channels, and across-channel comparisons of temporal envelope information have been suggested to contribute to CMR. The present study extends an earlier model that includes an equalization-cancellation (EC) stage for the processing of modulations across the audio-frequency channels by a non-linear peripheral filtering stage. In the framework of the model, the combination and interaction of three main mechanisms were assessed: (i) suppression, (ii) within-auditory-channel cues related to amplitude modulations, and (iii) across-auditory-channel processes at higher, retro-cochlear stages. Experiments are presented to examine the relative role of these mechanisms. In particular, the influence of level and effects of auditory grouping on CMR were investigated.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Piechowiak, T. (Intern)
Number of pages: 2
Pages: 179-180
Publication date: 2008

Host publication information
Title of host publication: Fortschritte der Akustik
Main Research Area: Technical/natural sciences
Conference: 34th German Annual Conference on Acoustics, Dresden, Germany, 10/03/2008 - 10/03/2008
comodulation masking release
Electronic versions:
daga2008_tobias.pdf
Source: orbit
Source-ID: 250598
Publication: Research › Article in proceedings – Annual report year: 2008
Open-plan office environments: A laboratory experiment to examine the effect of office noise and temperature on human perception, comfort and office work performance

General information
State: Published
Organisations: Section for Indoor Environment, Department of Civil Engineering, Hearing Systems, Department of Electrical Engineering
Authors: Balazova, I. (Ekstern), Clausen, G. (Intern), Rindel, J. H. (Ekstern), Poulsen, T. (Intern), Wyon, D. P. (Intern)
Pages: Paper ID: 703
Publication date: 2008

Host publication information
Title of host publication: the 11. International conference on Indoor Air Quality and Climate - Indoor Air 2008
Publisher: Technical University of Denmark (DTU)
Main Research Area: Technical/natural sciences
Conference: 11th International Conference on Indoor Air Quality and Climate, Copenhagen, Denmark, 17/08/2008 - 17/08/2008
Source: orbit
Source-ID: 232871
Publication: Research - peer-review › Article in proceedings – Annual report year: 2008

Pitch perception in hearing-impaired listeners

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Santurette, S. (Intern), Dau, T. (Intern)
Publication date: 2008

Publication information
Original language: English
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 253783
Publication: Research › Sound/Visual production (digital) – Annual report year: 2008

Questionnaire investigation of musicians’ use of hearing protectors, self reported hearing disorders and their experience of their working environment

Musicians in symphony orchestras are exposed to harmful sound levels. Although research shows that industrial workers have a higher propensity to noise induced hearing loss, musicians can also develop a hearing loss from noise exposure. Furthermore, musicians can suffer from tinnitus, hyperacusis, and distortion, among other hearing disorders, which can affect their work more severely than a hearing loss. This study investigated the use of hearing protectors, the prevalence of self-reported hearing disorders among musicians, and the importance of these hearing disorders to the musicians. The musicians at three Danish symphony orchestras were asked to complete a questionnaire on the topic. Results showed that Danish musicians are aware of the dangers of loud music, yet they rarely use hearing protectors and not always correctly; however, musicians with hearing disorders use hearing protectors more frequently. In addition, the musicians questioned suffered from different hearing disorders. Education is needed to change musicians’ opinion of hearing conservation and hearing protectors. The education must be directed to both the musicians and the administration of the symphony orchestras.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research, Heikki Helimäki Ltd
Authors: Laitinen, H. (Ekstern), Poulsen, T. (Intern)
Pages: 160-168
Publication date: 2008
Main Research Area: Technical/natural sciences

Publication information
Journal: International Journal of Audiology
Volume: 47
Issue number: 4
Relating the absence of binaural pitch percept to retro-cochlear impairment

General information
State: Published
The perceptual flow of phonetic feature processing

How does the brain process spoken language? It is our thesis that word intelligibility and consonant identification are insufficient by themselves to model how the speech signal is decoded - a finer-grained approach is required. In this study, listeners identified 11 different Danish consonants spoken in a Consonant + Vowel + [l] environment. Each syllable was processed so that only a portion of the original audio spectrum was present. Three-quarter-octave bands of speech, centered at 750, 1500, and 3000 Hz, were presented individually and in combination with each other. The conditional, posterior probabilities associated with phonetic-feature decoding were computed from confusion matrices in order to deduce the temporal flow of phonetic processing. Decoding the feature, Manner-of-Articulation, depends on accurate decoding of the feature Voicing (but not vice-versa), and decoding Place-of-Articulation requires precise decoding of Manner (but not the converse). From these data, we conclude that Voicing is processed prior to Manner-of-Articulation, and that Manner is decoded prior to Place-of-Articulation. Voicing and Manner cues are often correctly decoded in conditions where Place is not. This asymmetric pattern of feature decoding may provide extra-segmental information of utility for speech processing, particularly in adverse listening conditions.
**Bibliographical note**
Copyright (2008) Acoustical Society of America. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the Acoustical Society of America.

**Source**: orbit
**Source-ID**: 264274
**Publication**: Research - peer-review › Journal article – Annual report year: 2008

**Information from Impact Sounds: Normal and Impaired Hearing**

**General information**
- **State**: Published
- **Organisations**: Acoustic Technology, Department of Electrical Engineering, Hearing Systems
- **Authors**: Kirkwood, B. C. (Intern), Poulsen, T. (Intern), Naylor, G. (Ekstern)
- **Number of pages**: 262
- **Publication date**: Apr 2007

**Publication information**
- **Original language**: English
- **Main Research Area**: Technical/natural sciences
- **Electronic versions**:
Hearing aid processing of loud speech and noise signals: Consequences for loudness perception and listening comfort.

Hearing aid processing of loud speech and noise signals: Consequences for loudness perception and listening comfort. Sound processing in hearing aids is determined by the fitting rule. The fitting rule describes how the hearing aid should amplify speech and sounds in the surroundings, such that they become audible again for the hearing impaired person. The general goal is to place all sounds within the hearing aid users' audible range, such that speech intelligibility and listening comfort become as good as possible. Amplification strategies in hearing aids are in many cases based on empirical research -for example investigations of loudness perception in hearing impaired listeners. Most research has been focused on speech and sounds at medium input-levels (e.g., 60-65 dB SPL). It is well documented that for speech at conversational levels, hearing-aid users prefer the signal to be amplified by approximately half the amount of the hearing loss (in dB). This places the amplified speech signal approximately in the middle of the users' audible range, at a comfortable listening level. However, there has been little research on the optimal gain-prescription for soft and loud sounds. At present, such prescriptions are based mainly on logic, as there is limited evidence on what type of amplification is best for these input-levels. The focus of the PhD-project has been on hearing aid processing of loud speech and noise signals. Previous research, investigating the preferred listening levels for soft and loud sounds, has found that both normal-hearing and hearing-impaired listeners prefer loud sounds to be closer to the most comfortable loudness-level, than suggested by common non-linear fitting rules. During this project, two listening experiments were carried out. In the first experiment, hearing aid users listened to loud speech and noise signals with built-in level-variation (62 – 82 dB SPL). The signals had been compressed with seven different compression ratios, in the range from 1:1 to 10:1, yielding different degree of overall level-variation in the processed signals. Subjects rated the signals in regard to perceived level variation, loudness and overall acceptance. In the second experiment, two signals containing speech and noise at 75 dB SPL RMS-level, were compressed with six compression ratios from 1:1 to 10:1 and three release times from 40 ms to 4000 ms. In this experiment, subjects rated the signals in regard to loudness, speech clarity, noisiness and overall acceptance. Based on the results, a criterion for selecting compression parameters that yield some level-variation in the output signal, while still keeping the overall user-acceptance at a tolerable level, is suggested. It is also discussed how differences in speech and noise components seem to influence listeners ratings of the test signals. General recommendations for a fitting rule, that takes into account the spectral and temporal characteristics of the input signal, is given together with suggestions for further studies. Finally, the experimental methods used for the listening tests in this project are discussed «««.

General information
State: Published
Organisations: Acoustic Technology, Department of Electrical Engineering, Hearing Systems, Widex A/S
Authors: Schmidt, E. (Intern), Poulsen, T. (Intern), Ludvigsen, C. (Ekstern)
Number of pages: 198
Publication date: Jan 2007

Publication Information
ISBN (Print): 97-88-79118471-0
Original language: English
Main Research Area: Technical/natural sciences
Electronic versions:
PhD-report Erik (final version 10-1-07).pdf
Source: orbit
Source-ID: 195742
Publication: Research › Ph.D. thesis – Annual report year: 2007

Constrained ICA for the Analysis of High Stimulus Rate Auditory Evoked Potentials

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Harte, J. (Intern)
Pages: 609-616
Publication date: 2007

Host publication information
Title of host publication: 7th International Conference on Independent Component Analysis and Signal Separation
Place of publication: Berlin
Publisher: Springer
Main Research Area: Technical/natural sciences
Conference: Independent component analysis and signal separation, London, UK, 01/01/2007
**Linguistic Scene Analysis and the Importance of Synergy**

This chapter explores the possibility that speech is decoded using cross-spectral and cross-modal integration strategies that are inherently synergistic. Combining information from separate spectral channels or across modalities may result in far greater intelligibility and phonetic recognition than predicted by linear-integration models. This is because decoding speech relies on multi-tier processing strategies that are opportunistic and idiosyncratic. Models incorporating synergistic integration are more likely to predict linguistic comprehension than conventional, linear approaches, particularly in challenging listening conditions.

**Mechanisms of within- and across- channel processing in comodulation masking release**

**Training of Speechreading for Severely Hearing-Impaired Persons by Human and Computer**

This paper describes evaluation results for a software programme that is intended to be used as a training-aid for lipreading in German. Tests were carried out in schools for hearing-impaired children in Germany which indicate that the ability to lipread increases significantly already after use of the software during a short period of time.
**Distant teaching of chamber music via local area networks**

In this paper we present a study on teaching chamber music via internet. The application for this setup is for a highly reputed teacher to teach professional musicians at a very high level. Usually, all participants would have to fly from all over the world in order to work together. Therefore, it would be of great value, if these teaching lessons could be done via internet. Several audio and video devices and different audio setups have been tested. The results indicate that MPEG 2 broadcast devices with two microphones are suitable for this task.

**Melody recognition with binaural-pitch stimuli in normal-hearing and hearing-impaired listeners**

In this study, we examine the effect of binaural pitch stimuli on melody recognition in normal-hearing and hearing-impaired listeners.
Auditory object formation affects modulation perception
Most sounds in our environment, including speech, music, animal vocalizations and environmental noise, have fluctuations in intensity that are often highly correlated across different frequency regions. Because across-frequency modulation is so common, the ability to process such information is thought to be a powerful survival strategy in the natural world (Klump, 1996; Nelken et al., 1999). Coherent modulations in one sound can aid in the detection of another sound (Hall et al., 1984; Durlach, 1963). On the other hand, modulation in one frequency region can also impede the detection or discrimination of modulation in other frequency regions (Yost et al., 1989). Although the neural substrates for across-frequency modulation processing remain unclear, recent studies have concentrated on brainstem structures (Pressnitzer et al., 2001). In this study it is shown that sounds occurring after the target sound in time determine whether or not across-frequency modulation effects are observed. The results suggest that the binding of sound elements into coherent auditory objects precedes aspects of modulation analysis and imply a cortical locus involving integration times of several hundred milliseconds. In other words, the modulation analysis necessary for signal detection is performed on objects, rather than frequency channels. The results place strong constraints on the search for neural correlates of this important aspect of auditory processing, and on future models of spectrotemporal processing.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Piechowiak, T. (Intern)
Number of pages: 17
Pages: 35-52
Publication date: 2005

Host publication information
Title of host publication: Hearing Aid Fitting: Symposium on Hearing Aid Fitting
Publisher: The Danavox Jubilee Foundation
ISBN (Print): 87-982422-0-2
Main Research Area: Technical/natural sciences
Conference: 21st Danavox Symposium "Hearing Aid Fitting", Kolding, Denmark, 31/08/2005 - 31/08/2005
Auditory grouping
Electronic versions:
danavox2005.doc
Source: orbit
Source-ID: 250594
Publication: Research » Article in proceedings – Annual report year: 2005

Principles of modulation processing in monaural vs. binaural hearing

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Centre for Applied Hearing Research
Authors: Piechowiak, T. (Intern)
Number of pages: 2
Pages: 569-571
Publication date: 2005

Host publication information
Title of host publication: Fortschritte der Akustik
Main Research Area: Technical/natural sciences
Conference: 32nd German Annual Conference on Acoustics, München, Germany, 20/03/2006 - 20/03/2006
modulation detection
Electronic versions:
Principles of modulation processing in monaural vs.doc
Source: orbit
Source-ID: 250596
Publication: Research » Article in proceedings – Annual report year: 2005

A Model for the representation of Speech Signals in Normal and Impaired Ears
A model of human auditory periphery, ranging from the outer ear to the auditory nerve, was developed. The model consists of the following components: outer ear transfer function, middle ear transfer function, basilar membrane velocity, inner hair cell receptor potential, inner hair cell probability of neurotransmitter release and auditory nerve fibre refractoriness. The model builds on previously published models, however, parameters for basilar membrane velocity and inner hair cell probability of neurotransmitter release were successfully fitted to model data from psychophysical and physiological data for normal hearing and impaired hearing. The psychophysical data consisted in forward masking data.
from three studies. The "temporal window model" was tested and found to account for the data, except for low frequency stimulus. It was suggested that the temporal window should be frequency dependent. Impaired hearing was modelled as a combination of outer- and inner hair cell loss. The percentage of dead inner hair cells was calculated based on a new computational method relating auditory nerve fibre thresholds to behavioural thresholds. Finally, a model of the entire auditory nerve fibre population was proposed for normal and impaired hearing.

**General information**
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering
Authors: Christiansen, T. U. (Intern), Poulsen, T. (Intern)
Publication date: May 2004

**Publication information**
ISBN (Print): 87-91184-42-8
Original language: English
Main Research Area: Technical/natural sciences
Electronic versions:
180647_Orsted2004-Thomas Ulrich Christiansen-A Model for the representation of Speech Signals in Normal and Impaired Ears.pdf
Source: orbit
Source-ID: 180647
Publication: Research › Ph.D. thesis – Annual report year: 2004

**Benefit from amplification of high frequencies in hearing impaired: aspects of dead regions and auditory acclimatization**

**General information**
State: Published
Organisations: Department of Electrical Engineering, Hearing Systems
Authors: Vestergaard, M. D. (Intern), Poulsen, T. (Intern), Naylor, G. (Ekstern)
Publication date: Apr 2004

**Publication information**
ISBN (Print): 87-91184-32-0
Original language: English
Main Research Area: Technical/natural sciences
Electronic versions:
Orsted2004-MartinVestergaard-Benefit from amplification of high frequencies in hearing impaired aspects of dead regions and auditory acclimatization.pdf
Source: orbit
Source-ID: 180635
Publication: Research › Ph.D. thesis – Annual report year: 2004

**International hearing protector standardization**
Hearing protectors shall fulfill some minimum requirements to their performance. As hearing protector manufacturers sell the products all over the world, the testing and certification of hearing protectors has become an international issue. The ISO working group WG17 under the headlines Acoustics, Noise, produce hearing protector standards to be used at an international level. The presentation will cover the ongoing work in WG17, including the revision of existing standards (ISO 4869-1, ISO 4869-3), upcoming new standards (ISO 4869-7) and the plans and status for future standards (performance in impulse noise, protectors with active noise reduction). Furthermore, an overview of the present European standards (CEN) and the relation to American and Australian/New Zealand standards will be discussed.

**General information**
State: Published
Organisations: Hearing Systems, Department of Acoustic Technology
Authors: Poulsen, T. (Intern)
Pages: 2293-2293
Publication date: 2002
Main Research Area: Technical/natural sciences

**Publication information**
Volume: 112
Issue number: 5
ISSN (Print): 0001-4966
Ratings:

BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.83 SJR 0.749 SNIP 1.27
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.802 SNIP 1.437 CiteScore 1.77
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.788 SNIP 1.423 CiteScore 1.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.705 SNIP 1.966 CiteScore 2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.763 SNIP 1.622 CiteScore 1.75
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.695 SNIP 1.642 CiteScore 1.68
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.754 SNIP 1.528
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.783 SNIP 1.717
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 0.848 SNIP 1.633
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 0.865 SNIP 1.647
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 0.752 SNIP 1.559
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 0.954 SNIP 1.749
Web of Science (2005): Indexed yes
Scopus rating (2004): SJR 0.77 SNIP 1.787
Web of Science (2004): Indexed yes
Scopus rating (2003): SJR 0.882 SNIP 1.712
Web of Science (2003): Indexed yes
Scopus rating (2002): SJR 0.87 SNIP 1.501
Web of Science (2002): Indexed yes
Scopus rating (2001): SJR 0.719 SNIP 1.467
Web of Science (2001): Indexed yes
Scopus rating (2000): SJR 0.621 SNIP 1.411
Web of Science (2000): Indexed yes
Scopus rating (1999): SJR 0.591 SNIP 1.319

Original language: English
Electronic versions:
On the relations among temporal integration for loudness, loudness discrimination, and the form of the loudness function.

(A)

Temporal integration for loudness was measured as a function of level from 2 to 60 dB SL using 2-, 10-, 50-, and 250-ms tones at 5 kHz. The adaptive 2I,2AFC procedure converged at the level required to make the variable stimulus just louder than the fixed stimulus. Thus the data yield estimates of the levels required to make tones of different durations equally loud and of the just noticeable differences for loudness level. Results for four listeners with normal hearing show that the amount of temporal integration, defined as the level difference between equally loud short and long tones, varies markedly with level and is largest at moderate levels. The effect of level increases as the duration of the short stimulus decreases and is largest for comparisons between the 2- and 250-ms tones. The loudness-level jnds are also largest at moderate levels and, contrary to traditional jnds for the level of two equal-duration tones, they do not appear to depend on duration. The level dependence of temporal integration and the loudness jnds are consistent with a loudness function \( \log(\text{loudness}) \) versus SPL that is flatter at moderate levels than at low and high levels. [Work supported by NIH-NIDCD R01DC02241
Scopus rating (2004): SJR 0.77 SNIP 1.787
Web of Science (2004): Indexed yes
Scopus rating (2003): SJR 0.882 SNIP 1.712
Web of Science (2003): Indexed yes
Scopus rating (2002): SJR 0.87 SNIP 1.501
Web of Science (2002): Indexed yes
Scopus rating (2001): SJR 0.719 SNIP 1.467
Web of Science (2001): Indexed yes
Scopus rating (2000): SJR 0.621 SNIP 1.411
Web of Science (2000): Indexed yes
Scopus rating (1999): SJR 0.591 SNIP 1.319
Original language: English
Electronic versions:
Poulsen.pdf
DOIs:
10.1121/1.415616

Bibliographical note
Copyright (1996) Acoustical Society of America. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the Acoustical Society of America.
Source: orbit
Source-ID: 264159
Publication: Research - peer-review › Journal article – Annual report year: 1996

Hearing aid measurements with speech and noise signals
An increasing number of hearing aid types include one or more features which are intentionally non-linear. In such devices measurement of frequency response and distortion using sweep tone measurements are typically of little relevance. Five different non-linear hearing aid types were used to evaluate three different broad-band measuring methods. The results revealed that these methods were meaningful in estimating average frequency response obtained with a specific input signal, but none of the three methods used in the study was able to evaluate separately the effects of the most important signal modifications: memoryless non-linearity like peak clipping, time-varying gain from AGC and additive internal noise.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, GN ReSound A/S, Widex A/S
Authors: Dyrlund, O. (Ekstern), Ludvigsen, C. (Ekstern), Olofsson, Å. (Intern), Poulsen, T. (Intern)
Pages: 153-157
Publication date: 1994
Main Research Area: Technical/natural sciences

Publication information
Journal: Scandinavian Audiology
Volume: 23
Issue number: 3
ISSN (Print): 0105-0397
Ratings:
BFI (2008): BFI-level 1
Scopus rating (2004): SJR 0.703 SNIP 2.403
Scopus rating (2003): SJR 0.307 SNIP 1.075
Scopus rating (2002): SJR 0.312 SNIP 1.506
Scopus rating (2001): SJR 0.571 SNIP 0.849
Scopus rating (2000): SJR 0.388 SNIP 0.754
Scopus rating (1999): SJR 0.587 SNIP 1.075
Original language: English
Source: orbit
Source-ID: 264544
Publication: Research - peer-review › Journal article – Annual report year: 1994

Hearing Threshold and Equal Loudness Level Contours of 1/3-octave Noise Bands in a Diffuse Sound Field
Hearing threshold levels and equal loudness level contours of 1/3-octave noise bands at 40 phons and 60 phon were measured for 27 normal hearing listeners in an approximately diffuse sound field. The threshold data in the frequency
range 125 Hz to 1 kHz were 3-6 dB higher than the values given in ISO 226-1987. The equal loudness data in the frequency range 125 Hz to 500 Hz exceeded the ISO 226-1987 values by 5-10 dB.

General information
State: Published
Organisations: Hearing Systems, Department of Electrical Engineering, Technical University of Denmark
Authors: Nielsen, M. K. E. (Ekstern), Poulsen, T. (Intern)
Pages: 306-310
Publication date: 1994
Main Research Area: Technical/natural sciences

Publication information
Journal: Acustica
Volume: 80
Issue number: 3
ISSN (Print): 0001-7884
Ratings:
BFI (2008): BFI-level 1
Scopus rating (2004): SJR 0.585 SNIP 2.135
Scopus rating (2003): SJR 0.464 SNIP 1.995
Scopus rating (2002): SJR 0.358 SNIP 1.307
Scopus rating (2001): SJR 0.279 SNIP 0.533
Scopus rating (2000): SJR 0.418 SNIP 0.761
Scopus rating (1999): SJR 0.514 SNIP 1.017
Original language: English
Source: orbit
Source-ID: 259087
Publication: Research - peer-review › Journal article – Annual report year: 1994

Projects:

Clinical auditory profiling and hearing-aid fitting strategies
In audiological clinics, the choice of a hearing aid and the adjustment of its amplification and processing parameters are today mostly based on the audiogram, a measure of pure-tone hearing sensitivity at different frequencies. While adjusting the gain of a hearing aid based on the loss of sensitivity reflected by the audiogram can be successful in restoring audibility of soft sounds and improving speech intelligibility in quiet situations, it is well established that hearing-impaired listeners still experience difficulty with understanding speech in more complex listening situations that are typical of everyday life, such as noisy and reverberant environments (Moore, 2007). Despite amplification from the hearing aid, sounds are thus still perceived as distorted, and this “distortion loss” (Plomp, 1978) is still a challenge to compensate for in practice.

The idea of the present project is to improve the hearing-aid fitting process and suggest parameter adjustment rationales based on a more complete evaluation of each patient’s hearing profile that reflects distortion loss as well. It is hypothesized that hearing-aid benefit can be improved by directly relating outcomes from such an extended clinical hearing profile to the choice of hearing-aid fitting.

Department of Electrical Engineering
Hearing Systems
Period: 01/07/2016 → 01/07/2019
Number of participants: 1
Audiology, hearing aid, hearing science
Number of related Ph.D. students: 1
Project participant:
Sanchez Lopez, Raul (Intern)
Project

Spatial release from masking in complex acoustical scenes and the effect of hearing aid processing
Department of Electrical Engineering
Hearing Systems
Period: 04/01/2016 → 04/07/2016
Number of participants: 5
Speech intelligibility, Spatial Hearing, Spatial Release from Masking
Project participant:
Löw, Vera (Ekstern)
Supervisor:
Westermann, Adam (Ekstern)
Marschall, Marton (Intern)
Cubick, Jens (Intern)
Main Supervisor:
Dau, Torsten (Intern)

Project

The influence of visual cues on sound externalization
Department of Electrical Engineering
Hearing Systems
Period: 26/01/2015 → 26/07/2015
Number of participants: 4
Project participant:
Gil Carvajal, Juan Camilo (Ekstern)
Supervisor:
Cubick, Jens (Intern)
Santurette, Sébastien (Intern)
Dau, Torsten (Intern)

Project

Comparison of binaural microphones for externalization of sound
Department of Electrical Engineering
Hearing Systems
Period: 08/01/2015 → 05/07/2015
Number of participants: 4
Number of related Ph.D. students: 1
Project participant:
Sánchez Rodríguez, Cristian (Ekstern)
Supervisor:
Cubick, Jens (Intern)
Song, Wookeun (Ekstern)
MacDonald, Ewen (Intern)

Project

Evaluation of a clinical auditory profile in hearing-aid candidates
Department of Electrical Engineering
Hearing Systems
University of Copenhagen
Rigshospitalet
Bispebjerg University Hospital
Period: 03/02/2014 → 19/12/2014
Number of participants: 3
Project participant:
Santurette, Sébastien (Intern)
Thorup, Nicoline (Ekstern)
Supervisor:
Friis, Morten (Ekstern)
Activities:

41st MidWinter Meeting of the Association for Research in Otolaryngology
Period: 9 Feb 2018 → 14 Feb 2018
Andreu Paredes Gallardo (Participant)
Department of Electrical Engineering
Hearing Systems

Description
Poster presentation at the 41st MidWinter Meeting of the Association for Research in Otolaryngology
Degree of recognition: International
Links:
http://www.aro.org/

Related event
41st MidWinter Meeting of the Association for Research in Otolaryngology
09/02/2018 → 14/02/2018
San Diego, United States
Activity: Attending an event › Participating in or organising a conference

A correlation metric in the envelope power spectrum domain for speech intelligibility prediction
Period: 2017
Helia Relano Iborra (Guest lecturer)
Department of Electrical Engineering
Hearing Systems

Description
A powerful tool to investigate speech perception is the use of speech intelligibility prediction models. Recently, a model was presented, termed correlation-based speech-based envelope power spectrum model (sEPSMcorr), that uses a correlation-based back end at the output of an audio-frequency and modulation-frequency selective auditory preprocessing (Relaño-Iborra et al., 2016). The use of the correlation back-end extended the predictive power of earlier versions of the sEPSM framework (e.g. Jørgensen et al. 2013) towards conditions of non-linear signal processing, such as phase jitter and ideal binary mask processing. Moreover, the model was shown to account for conditions with fluctuating interferers, unlike other correlation-based models.

Here, the back end of the sEPSMcorr was combined with a more realistic auditory pre-processing front end adopted from the computational auditory signal processing and perception model (CASP; Jepsen et al., 2008). The preprocessing contains outer- and middle-ear filtering and a non-linear auditory filterbank (DRNL, López-Poveda and Meddis, 2001), followed by inner hair-cell transduction, adaptation and a modulation filterbank.

The predictions were compared to measured data in conditions of additive masking noise, phase jitter distortions, reverberation and noise-reduction algorithms. The effects of the back end as well as the different preprocessing stages on the predicted results were analyzed. The modelling framework could be useful for the design and evaluation of, e.g. speech transmission algorithms or hearing-instrument algorithms.

Documents:
spin_helia_final_v2

Related event
9th Speech in Noise Workshop
05/01/2017 → 06/01/2017
Oldenburg, Germany
Activity: Talks and presentations › Talks and presentations in private or public companies and organisations
A correlation metric in the envelope power spectrum domain for speech intelligibility prediction

Period: 2017

Helia Relano Iborra (Guest lecturer)
Department of Electrical Engineering
Hearing Systems

Description
A speech intelligibility model, named sEPSMcorr, is presented, which uses a modulation-frequency selective processing based on the (multi-resolution) speech-based envelope power spectrum model (mr-sEPSM; Jørgensen et al. 2013) in combination with a cross-correlation based back end inspired by the short-time objective intelligibility measure (STOI; Taal et al., 2011). The model can accurately predict data obtained with normal-hearing (NH) listeners for a broad range of listening conditions, including effects of stationary and fluctuating additive interferers as well as effects of non-linear distortions, such as spectral subtraction, phase jitter and ideal binary mask (IBM) processing. The model has a larger predictive power than both the original mr-sEPSM (which fails in the phase-jitter and IBM conditions) and STOI (which fails to predict the influence of fluctuating interferers).

However, the sEPSMcorr preprocessing does not provide a flexible framework to predict individual speech intelligibility data from hearing impaired listeners. Thus, the back end of the sEPSMcorr was combined with a more realistic auditory pre-processing front end adapted from the computational auditory signal processing and perception model (CASP; Jepsen et al., 2008). The preprocessing contains outer- and middle-ear filtering and a non-linear auditory filterbank (DRNL, López-Poveda and Meddis, 2001), followed by inner hair-cell transduction, adaptation and a modulation filterbank.

The predictions of the sEPSM-based and the CASP-based models were compared with respect to measured data (NH) in conditions of additive masking noise, phase jitter distortions, reverberation and noise-reduction algorithms. The effects of the back end as well as the different preprocessing stages on the predicted results were analyzed. The resulting modelling framework could be useful for the design and evaluation of, e.g. speech transmission algorithms or hearing-instrument algorithms.

Documents:
ARCHES_poster_final3

Related event
ARCHES/ICANHEAR 2016: Audiological Research Cores in Europe (ARCHES) meeting and Improved Communication through Applied Hearing Research (ICanHear) conference
Zurich, Switzerland
Activity: Talks and presentations › Conference presentations

Extending a computational model of auditory processing towards speech intelligibility prediction

Period: 2017

Helia Relano Iborra (Guest lecturer)
Department of Electrical Engineering
Hearing Systems

Description
A speech intelligibility model is presented, based on the computational auditory signal processing and perception model (CASP; Jepsen et al., 2008). CASP has previously been shown to successfully predict psychoacoustic data of normal hearing (NH) listeners obtained in conditions of, e.g., spectral masking, amplitude-modulation detection, and forward masking (Jepsen et al., 2008). Furthermore, CASP can be tuned to model data from individual hearing-impaired listeners in different behavioral experiments (Jepsen and Dau, 2011). In this study, the CASP model is investigated as a predictor of intelligibility for Danish sentences for NH listeners.

The model receives the clean and degraded speech as input. The signals are processed through outer- and middle-ear filtering, a non-linear auditory filterbank (DRNL, López-Poveda and Meddis, 2001), adaptation loops, and a modulation filterbank. The internal representations produced at the end of these stages are analyzed using a correlation-based back end.

Here, predictions of speech intelligibility obtained with the speech-based CASP implementation are presented and compared to speech intelligibility data measured in conditions of additive noise, phase jitter, spectral subtraction, ideal binary mask processing and reverberation.

Related event
International Symposium on Auditory and Audiological Research
23/08/2017 → 25/08/2017
Nyborg, Denmark
Activity: Talks and presentations › Conference presentations
**Proceedings of the International Symposium on Auditory and Audiological Research (Journal)**

Period: 2017 → …

Sébastien Santurette (Editor)
Torsten Dau (Editor)

Department of Electrical Engineering
Hearing Systems

**Description**
Proceedings of ISAAR: International Symposium on Auditory and Audiological Research
Degree of recognition: International

Links:
https://proceedings.isaar.eu

**Related journal**

Proceedings of the International Symposium on Auditory and Audiological Research

**An extended test battery for characterizing hearing deficits**

Period: 29 Sep 2017

Raul Sanchez Lopez (Speaker)
Federica Bianchi (Other)
Michal Fereczkowski (Other)
Sébastien Santurette (Other)
Torsten Dau (Other)

Department of Electrical Engineering
Hearing Systems

**Related event**

Dansk Teknisk Audiologisk Selskab årsmøde 2017
29/09/2017 → 30/09/2017
Activity: Talks and presentations › Conference presentations

**University of Salamanca**

Period: 1 Sep 2017 → 31 Dec 2017

Helia Relano Iborra (Visiting researcher)

Department of Electrical Engineering
Hearing Systems

**Description**
4 months research stay at the Auditory Computation & Psychoacoustics group of the Institute of Neurosciences f the UNiversity of Salamanca with Professor Enrique A. Lopez-Poveda
Activity: Visiting an external institution › Visiting another research institution

**Trends in Hearing (Journal)**

Period: Aug 2017

Helia Relano Iborra (Reviewer)

Department of Electrical Engineering
Hearing Systems

**Related journal**

Trends in Hearing
International Symposium on Auditory and Audiological Research

*Period: 23 Aug 2017 → 25 Aug 2017*

Andreu Paredes Gallardo (Participant)

Department of Electrical Engineering

**Hearing Systems**

**Description**

Oral presentation at the International Symposium on Auditory and Audiological Research

*Degree of recognition: International*

**Related event**

International Symposium on Auditory and Audiological Research

*23/08/2017 → 25/08/2017*

Nyborg, Denmark

Activity: Attending an event › Participating in or organising a conference

**Auditory profiling through computational data analysis**

*Period: 19 Aug 2017*

Raul Sanchez Lopez (Other)

Federica Bianchi (Other)

Michal Fereczkowski (Other)

Sébastien Santurette (Other)

Torsten Dau (Other)

Department of Electrical Engineering

**Hearing Systems**

**Description**

Nowadays, the pure-tone audiogram is the main tool used to characterize the degree of hearing loss and to fit hearing aids. However, the perceptual consequences of a hearing loss are typically associated not only with a loss of sensitivity, but also with a loss of clarity (distortion loss) that is not captured by the audiogram. Detailed characterization of hearing deficits can be complex and it has to be simplified in order to efficiently investigate the specific compensation needs of individual listeners. The aim of this study is to characterize individual hearing deficits by means of a test battery that allows to capture the diverse aspects of hearing loss, considering not only the loss of sensitivity but also supra-threshold distortions.

It was hypothesized that any listeners’ hearing can be characterized along two dimensions: distortion type I and distortion type II. While distortion type I can be linked to factors affecting audibility, distortion type II is considered as a non-audibility-related distortion, or clarity loss. To evaluate our hypothesis, the data from two studies was re-analyzed using a data-driven approach. Both studies carried out an extensive battery of psychoacoustic tests on potential hearing-aid users. The new analysis was based on an archetypal analysis and uses unsupervised learning to identify extreme patterns in the data which provide the basis for different auditory profiles. Subsequently, a decision tree was obtained that enables a simple classification of the listeners into one of the profiles.

This novel approach provided evidence for the existence of four different “auditory profiles” in the data. The most significant predictors for the profile identification were related to temporal processing, peripheral compression, and speech perception. The current approach is promising for identifying the most relevant tests for auditory profiling and considering new fitting strategies based on the individual’s deficits.

**Related event**


*19/08/2017 → 19/08/2017*

Stockholm, Sweden

Activity: Talks and presentations › Conference presentations
Conference on Implantable Auditory Prosthesis  
Period: 16 Jul 2017 → 21 Jul 2017  
Andreu Paredes Gallardo (Participant)  
Department of Electrical Engineering  
Hearing Systems

Description
Poster presentation at the conference

Links:
http://ciaphome.org/index.html (Home site)

Related event
Conference on Implantable Auditory Prosthesis  
16/07/2017 → 21/07/2017  
Lake Tahoe, United States  
Activity: Attending an event › Participating in or organising a conference

The speech-based envelope power spectrum model (sEPSM) family: Development, achievements, and current challenges  
Period: 29 Jun 2017  
Helia Relano Iborra (Guest lecturer)  
Department of Electrical Engineering  
Hearing Systems

Description
Intelligibility models provide insights regarding the effects of target speech characteristics, transmission channels and/or auditory processing on the speech perception performance of listeners. In 2011, Jørgensen and Dau proposed the speech-based envelope power spectrum model [sEPSM, Jørgensen and Dau (2011). J. Acoust. Soc. Am. 130(3), 1475-1487]. It uses the signal-to-noise ratio in the modulation domain (SNRenv) as a decision metric and was shown to accurately predict the intelligibility of processed noisy speech. The sEPSM concept has since been applied in various subsequent models, which have extended the predictive power of the original model to a broad range of conditions. This contribution presents the most recent developments within the sEPSM "family:" (i) A binaural extension, the B-sEPSM [Chabot-Leclerc et al. (2016). J. Acoust. Soc. Am. 140(1), 192-205] which combines better-ear and binaural unmasking processes and accounts for a large variety of spatial phenomena in speech perception; (ii) a correlation-based version [Relaño-Iborra et al. (2016). J. Acoust. Soc. Am. 140(4), 2670-2679] which extends the predictions of the early model to non-linear distortions, such as phase jitter and binary mask-processing; and (iii) a recent physiologically inspired extension, which allows to functionally account for effects of individual hearing impairment on speech perception.  
Degree of recognition: International  
Links:
http://dx.doi.org/10.1121/1.4989047

Related event
173rd Meeting of the Acoustical Society of America and the 8th Forum Acusticum  
25/06/2017 → 29/06/2017  
Boston , United States  
Activity: Talks and presentations › Conference presentations

Listening to music with a cochlear implant: Limitations and possible solutions  
Period: 8 Mar 2017  
Jeremy Marozeau (Invited speaker)  
Sébastien Santurette (Invited speaker)  
Department of Electrical Engineering  
Hearing Systems

Description
Although the cochlear implant can restore the perception of speech in quiet environments remarkably well, CI users are still facing many challenges in order to perceive music. In this talk, we describe how musical dimensions (pitch, tempo, timbre,...) are affected by the sound processor and a few solutions that could be used to improve the enjoyment of music by CI users.
Although the cochlear implant can restore the perception of speech in quiet environments remarkably well, CI users are still facing many challenges in order to perceive music. In this talk, we describe how musical dimensions (pitch, tempo, timbre,...) are affected by the sound processor and a few solutions that could be used to improve the enjoyment of music by CI users.

Related event

Nordiske Konference - Hørelse, kognition, kommunikation
18/03/2015 → …
Fredericia, Denmark
Activity: Talks and presentations › Conference presentations

Nordiske Konference - Hørelse, kognition, kommunikation
Period: 8 Mar 2017
Wiebke Lamping (Participant)
Steffen Spangmose Pedersen (Participant)
Department of Electrical Engineering
Hearing Systems

Description
Listening to music with a cochlear implant: Limitations and possible solutions

Although the cochlear implant can restore the perception of speech in quiet environments remarkably well, CI users are still facing many challenges in order to perceive music. In this talk, we describe how musical dimensions (pitch, tempo, timbre,...) are affected by the sound processor and a few solutions that could be used to improve the enjoyment of music by CI users.

Description
Visiting Cochlear House, Melbourne during my external research stay
Degree of recognition: International
Activity: Visiting an external institution › Visiting another research institution

ARCHES/ICANHEAR 2016
Andreu Paredes Gallardo (Participant)
Department of Electrical Engineering
Hearing Systems

Description
Poster contribution

Related event

ARCHES/ICANHEAR 2016: Audiological Research Cores in Europe (ARCHES) meeting and Improved Communication through Applied Hearing Research (ICanHear) conference
Zurich, Switzerland
Activity: Attending an event › Participating in or organising a conference

**Music and CI symposium**
Andreu Paredes Gallardo (Participant)
Department of Electrical Engineering
Hearing Systems

**Description**
Poster contribution

Poster contribution to the first international symposium on music and cochlear implants

**Related event**
**Music and CI symposium**
13/10/2016 → 14/10/2016
Snekkersten, Denmark
Activity: Attending an event › Participating in or organising a conference

**Beyond the audiogram: How can we achieve better hearing rehabilitation?**
*Period: 9 Sep 2016*
Sébastien Santurette (Invited speaker)
Department of Electrical Engineering
Hearing Systems

**Links:**
http://dtas.dk/aarsmoede_2016.php

**Related event**
**Dansk Teknisk Audiologisk Selskab**
09/09/2016 → 10/09/2016
Stouby, Denmark
Activity: Talks and presentations › Conference presentations

**Dansk Teknisk Audiologisk Selskab**
*Period: 9 Sep 2016*
Raul Sanchez Lopez (Participant)
Department of Electrical Engineering
Hearing Systems

**Description**
Poster presentation: Spectro-temporal modulation sensitivity and discrimination in normal hearing and hearing-impaired listeners

Dansk Teknisk Audiologisk Selskab (DTAS) årsmøde 2016

**Related event**
**Dansk Teknisk Audiologisk Selskab**
09/09/2016 → 10/09/2016
Stouby, Denmark
Activity: Attending an event › Participating in or organising workshops, courses, seminars etc.

**Effects of musical experience on pitch discrimination**
*Period: 9 Sep 2016*
Sébastien Santurette (Invited speaker)
Department of Electrical Engineering
Hearing Systems
Links:
http://dtas.dk/aarsmoede_2016.php

Related event

Dansk Teknisk Audiologisk Selskab
09/09/2016 → 10/09/2016
Stouby, Denmark
Activity: Talks and presentations › Conference presentations

Interspeech 2016
Period: 8 Sep 2016 → 12 Sep 2016
Thomas Bentsen (Participant)
Department of Electrical Engineering
Hearing Systems

Description
Attending and presenting a poster at the 2016 conference.
Degree of recognition: International

Related event

Interspeech 2016
08/09/2016 → 12/09/2016
San Francisco, Ca, United States
Activity: Attending an event › Participating in or organising a conference

Poster presentation
Period: 8 Sep 2016 → 12 Sep 2016
Thomas Bentsen (Speaker)
Department of Electrical Engineering
Hearing Systems

Description
Poster presentation
Degree of recognition: International

Related event

Interspeech 2016
08/09/2016 → 12/09/2016
San Francisco, Ca, United States
Activity: Talks and presentations › Conference presentations

Academic communication and effective presenting
Andreu Paredes Gallardo (Participant)
Department of Electrical Engineering
Hearing Systems

Description
Participated on the course

Participation in the course "Academic communication and effective presenting", part of the Tampere University Summer School program.
Related event

**Academic communication and effective presenting: University of Tampere Summer School**
08/08/2016 → 16/08/2016
Tampere, Finland
Activity: Attending an event › Participating in or organising workshops, courses, seminars etc.

**Improving Cochlear Implant Performance**
Period: 1 Jul 2016
Andreu Paredes Gallardo (Participant)
Department of Electrical Engineering

**Description**
Attended the talks and contributed with a poster presentation

Participation in the yearly meeting "Improving Cochlear Implant Performance" organized by UCL (London)

Related event

**Workshop on Improving Cochlear Implant Performance**
01/07/2016 → 01/07/2016
London, United Kingdom
Activity: Attending an event › Participating in or organising workshops, courses, seminars etc.

**International Symposium on Objective Measures in Auditory Implants**
Andreu Paredes Gallardo (Speaker)
Department of Electrical Engineering

**Description**
Oral presentation at the 9th International Symposium on Objective Measures in Auditory Implants (Szeged, Hungary)
Degree of recognition: International
Documents:
OMAI-2016-Abstract-Book
Links:
http://objectivemeasures.org/ (OMAI website)

Related event

**International Symposium on Objective Measures in Auditory Implants**
15/06/2016 → 18/06/2016
Szeged, Hungary
Activity: Talks and presentations › Conference presentations

**Predicting speech intelligibility based on a correlation metric in the modulation power domain**
Period: 5 Apr 2016
Helia Relano Iborra (Guest lecturer)
Department of Electrical Engineering

**Related external organisation**

**Eriksholm Research Centre**
Denmark
Activity: Talks and presentations › Talks and presentations in private or public companies and organisations
Teaching and learning DTU
Andreu Paredes Gallardo (Participant)
Department of Electrical Engineering
Hearing Systems

Description
Participation in the course Teaching and Learning organized by the DTU learning lab

Related event

Teaching and learning DTU
08/04/2014 → 11/04/2014
Kgs. Lyngby, Denmark
Activity: Attending an event › Participating in or organising workshops, courses, seminars etc.

The Danavox Jubilee Foundation (Publisher)
Period: 2015 → 2016
Sébastien Santurette (Editor)
Department of Electrical Engineering
Hearing Systems

Description
Proceedings of ISAAR: International Symposium on Auditory and Audiological Research
5th ISAAR: Individual Hearing Loss - Characterization, Modelling, Compensation Strategies

Related Publisher
The Danavox Jubilee Foundation
Local database
Activity: Research › Series editor

EP Master class IV: Cortical Evoked Potentials
Period: 3 Dec 2015
Andreu Paredes Gallardo (Participant)
Department of Electrical Engineering
Hearing Systems

Description
Introduction to electrophysiological recordings with CI users, theoretical concepts and practice sessions.
Participation in a one day course at Cochlear Academy (Mechelen, Belgium)

Related event

EP Master class IV: Cortical Evoked Potentials
03/12/2015 → 03/12/2015
Mechelen, Belgium
Activity: Attending an event › Participating in or organising workshops, courses, seminars etc.

Talk at Audiological Research Cores in Europe
Period: 17 Nov 2015
Thomas Bentsen (Speaker)
Department of Electrical Engineering
Hearing Systems
Description
Talk at Audiological Research Cores in Europe (ARCHES)
Degree of recognition: International

Related event

Audiological Research Cores in Europe 2015
16/11/2015 → 17/11/2015
Groningen, Netherlands
Activity: Talks and presentations › Conference presentations

INTERSPERCE 2015
Period: 6 Sep 2015 → 10 Sep 2015
Thomas Bentsen (Participant)
Department of Electrical Engineering
Hearing Systems

Description
Attending the Interspeech 2015 conference
Degree of recognition: International

Related event

INTERSPERCE 2015 : Speech beyond Speech
06/09/2015 → 10/09/2015
Dresden, Germany
Activity: Attending an event › Participating in or organising a conference

Analysis of correlated data: Mixed Linear Models
Period: 1 Sep 2015 → 25 Dec 2015
Andreu Paredes Gallardo (Participant)
Department of Electrical Engineering
Hearing Systems

Description
Analysis of correlated data: Mixed Linear Models

Related event

Analysis of correlated data: Mixed Linear Models
31/08/2015 → 25/12/2015
Kongens Lyngby, Denmark
Activity: Attending an event › Participating in or organising workshops, courses, seminars etc.

Communicating Advanced Topics in Electrical Engineering
Period: 1 Sep 2015 → 25 Dec 2015
Andreu Paredes Gallardo (Participant)
Department of Electrical Engineering
Hearing Systems

Description
Communicating Advanced Topics in Electrical Engineering

Related event

Communicating Advanced Topics in Electrical Engineering: Course Number 31920
31/08/2015 → 25/12/2015
Denmark
Activity: Attending an event › Participating in or organising workshops, courses, seminars etc.
Evaluation of peripheral compression and auditory nerve fiber intensity coding using Auditory Steady-State Responses (ASSR)

Gerard Encina Llamas (Speaker)
Department of Electrical Engineering

Description
The compressive nonlinearity of the auditory system is assumed to be an epiphenomenon of a healthy cochlea and particularly, outer-hair cell function. Auditory steady-state responses (ASSR) reflects coding of the stimulus envelope. Recent research in animals shows that noise over-exposure, producing temporary threshold shifts, can cause auditory nerve fiber (ANF) deafferentation in predominantly low-spontaneous rate (SR) fibers. It is hypothesized here that deafferentation of low-SR fibers can lead to a reduction of ASSR amplitude at supra-threshold levels. ASSR input/output (I/O) functions were measured in two groups of normal-hearing adults at stimulus levels ranging from 20 to 90 dB SPL. First, multi-frequency ASSR I/O functions were obtained using a modulation depth of 85%. Secondly, ASSR were obtained using a single sinusoidally amplitude modulated (SAM) tone at four modulation depths (25, 50, 85 and 100%). Results showed that ASSR growth functions exhibit compression of about 0.25 dB/dB. The slope for levels above 60 dB SPL showed more variability across subjects. The slope of ASSR I/O functions could be used to estimate peripheral compression simultaneously at four frequencies below 60 dB SPL, while the slope above 60 dB SPL might be used to evaluate the integrity of intensity coding of low-SR fibers.

Documents:
Encina-Llamas_ISAAR 2015

Links:
http://dx.doi.org/10.13140/RG.2.1.4485.1924

Related event
5th International Symposium on Auditory and Audiological Research
26/08/2015 → 28/08/2015
Nyborg, Denmark
Activity: Talks and presentations › Conference presentations

5th International Symposium on Auditory and Audiological Research
Andreu Paredes Gallardo (Participant)
Department of Electrical Engineering

Description
Presentation of a poster at the International Symposium on Auditory and Audiological Research, At Nyborg, Denmark.

Related event
5th International Symposium on Auditory and Audiological Research
26/08/2015 → 28/08/2015
Nyborg, Denmark
Activity: Attending an event › Participating in or organising workshops, courses, seminars etc.

The Analysis of Sensory and Consumer Data
Andreu Paredes Gallardo (Participant)
Department of Electrical Engineering

Related event
The Analysis of Sensory and Consumer Data: 02930
17/08/2015 → 21/08/2015
**Evaluation of cochlear processing and auditory nerve fiber intensity coding using Auditory Steady-State Responses (ASSR)**

**Period:** 11 May 2015

**Gerard Encina Llamas (Speaker)**

**Department of Electrical Engineering**

**Hearing Systems**

**Description**

**Objectives:**
The compressive nonlinearity of the peripheral auditory system is commonly assumed to be a result of healthy outer-hair cell function, and to be a good indicator of the system's integrity. It has recently been shown that auditory steady-state responses (ASSR) elicited by sinusoidally amplitude modulated (SAM) tones, with modulation frequencies around 80 – 100 Hz, show compressive growth as a function of stimulus intensity for medium levels. These responses are thought to reflect coding of the acoustical stimulus envelope at the level of the brainstem, hence after cochlear processing. Recent research in laboratory animals shows that noise exposure producing temporary threshold shifts can cause auditory nerve fiber (ANF) deafferentation, predominantly affecting low-spontaneous rate (SR) fibers. In the present study, it is hypothesized that deafferentation of low SR fibers leads to a reduction of ASSR amplitude at levels coded by this class of fibers.

**Methods:**
Multi-channel ASSR input/output (I/O) functions were measured in two groups of audiometrically normal-hearing adults. For the first group, ASSR I/O functions were obtained at stimulation sound pressure levels (SPL) ranging from 20 to 80 dB in steps of 5 dB using multi-frequency stimulation with four octave-spaced SAM tones and a fixed modulation depth of 85%. For the second group, ASSR I/O functions were obtained using a single SAM tone presented at 3 levels below 60 dB SPL and at 9 levels in the range from 60 to 90 dB SPL. ASSR growth functions were recorded at four modulation depths (65, 75, 85 and 100%).

**Results:**
Results showed that ASSR growth functions can be measured in NH listeners for input levels between 20 and 80-90 dB SPL. Significant compression was found for input levels between 30 to 60 dB SPL, and a saturation of the growth functions was observed at levels above 60 dB SPL. The lower-level part of the ASSR I/O functions showed compression of about ~0.25 dB/db, which is similar to compression recorded directly in animal cochleae. The slope for levels above 60 dB SPL showed larger variability across subjects.

**Conclusions:**
The slope of ASSR I/O functions could be used to estimate peripheral compression simultaneously at four frequencies. First results suggest, that the slope of ASSR I/O functions above 60 dB SPL might be used to evaluate the integrity of intensity coding of low-SR fibers. Hence, ASSR might prove to be useful for evaluation of both cochlear and ANF integrity.

**Funding**
This work was funded by the Oticon Center of Excellence for Hearing and Speech Sciences at the Technical University of Denmark.

**Documents:**
[Encina-Llamas_IERASG_2015](#)

**Links:**

http://dx.doi.org/10.13140/RG.2.1.2912.3280

**Related event**

**XXIV Biennial Symposium of the International Evoked Response Audiology Study Group**

10/05/2015 → 14/05/2015

Busan, Korea, Republic of

Activity: Talks and presentations › Conference presentations

---

**Improving Performance with Cochlear Implants**

Period: 21 Apr 2015

Andreu Paredes Gallardo (Participant)

Department of Electrical Engineering

Hearing Systems

**Related event**

**Improving Performance with Cochlear Implants**

21/04/2015 → 21/04/2015

London, United Kingdom

Activity: Attending an event › Participating in or organising workshops, courses, seminars etc.

---

**Fra øre til hjerne: tværfaglig forskning i hørelse og kognition**

Period: 18 Mar 2015

Sébastien Santurette (Invited speaker)

Department of Electrical Engineering

Hearing Systems

**Related event**

**Nordiske Konference - Hørelse, kognition, kommunikation**

18/03/2015 → …

Fredericia, Denmark

Activity: Talks and presentations › Conference presentations

---

**7th Workshop on Speech in Noise**

Period: 8 Jan 2015 → 9 Jan 2015

Sébastien Santurette (Organizer)

Department of Electrical Engineering

Hearing Systems

Links:

http://www.spin2015.dk

**Related event**

**7th Workshop on Speech in Noise**

08/01/2015 → 09/01/2015

Copenhagen, Denmark

Activity: Attending an event › Participating in or organising a conference

---

**7th Workshop on Speech in Noise**

Period: 8 Jan 2015 → 9 Jan 2015

Thomas Bentsen (Participant)

Department of Electrical Engineering

Hearing Systems
Related event

7th Workshop on Speech in Noise
08/01/2015 → 09/01/2015
Copenhagen, Denmark
Activity: Attending an event › Participating in or organising a conference

5th International Symposium on Auditory and Audiological Research
Period: 2014 → 2015
Sébastien Santurette (Organizer)
Department of Electrical Engineering
Hearing Systems

Description
Scientific committee - Abstract, programme, and manuscript coordinator - Webmaster
Links:
http://www.isaar.eu

Related event

5th International Symposium on Auditory and Audiological Research
26/08/2015 → 28/08/2015
Nyborg, Denmark
Activity: Attending an event › Participating in or organising a conference

Concepts and computational models of robust bottom-up signal encoding
Period: 30 Jan 2014 → 1 Feb 2014
Sébastien Santurette (Organizer)
Department of Electrical Engineering
Hearing Systems

Description
INSPIRE Winter School on concepts and computational models of robust bottom-up signal encoding
Links:
http://www.inspire-itn.eu

Related event

Concepts and computational models of robust bottom-up signal encoding
29/01/2014 → 01/02/2014
Copenhagen, Denmark
Activity: Attending an event › Participating in or organising a conference

The Danavox Jubilee Foundation (Publisher)
Period: 2013 → 2014
Sébastien Santurette (Editor)
Department of Electrical Engineering
Hearing Systems

Description
Proceedings of ISAAR: International Symposium on Auditory and Audiological Research
4th ISAAR: Auditory Plasticity - Listening with the Brain

Related Publisher

The Danavox Jubilee Foundation
Local database
Activity: Research › Series editor
21st International Congress on Acoustics
Period: 2 Jun 2013 → 7 Jun 2013
Nicolas Le Goff (Speaker)
Department of Electrical Engineering
Hearing Systems
Description
Poster presentation on "Spectral Integration of Interaural Time Differences in Auditory Localization"
Poster presentation
Documents:
SPECTRAL INTEGRATION OF INTERAURAL TIME DIFFERENCES IN AUDITORY LOCALIZATION

Related event
21st International Congress on Acoustics
02/06/2013 → 07/06/2013
Montreal, Canada
Activity: Talks and presentations › Conference presentations

36th Annual Midwinter Meeting of the Association for Research in Otolaryngology
Period: 16 Feb 2013 → 20 Feb 2013
Nicolas Le Goff (Participant)
Department of Electrical Engineering
Hearing Systems
Description
Abstract submission and Poster presentation
Documents:
Abstract
Poster

Related event
36th Annual Midwinter Meeting of the Association for Research in Otolaryngology
16/02/2013 → 16/02/2013
Baltimore, MD, United States
Activity: Attending an event › Participating in or organising a conference

International Symposium on Auditory and Audiological Research
Period: 2012 → 2013
Sébastien Santurette (Organizer)
Department of Electrical Engineering
Hearing Systems
Description
Abstract, programme, and manuscript coordinator - Webmaster
Abstract, programme, and manuscript coordinator - Webmaster; International Symposium on Auditory and Audiological Research (ISAAR) 2013

Related event
International Symposium on Auditory and Audiological Research
28/08/2013 → 30/08/2013
Nyborg, Denmark
Activity: Attending an event › Participating in or organising a conference
6th Annual Meeting of the Audiological Research Cores in Europe
Period: Nov 2012
Sébastien Santurette (Organizer)
Department of Electrical Engineering
Hearing Systems

Description
6th Meeting of European ARCHES network
Audiological Research Cores in Europe (ARCHES)
Links:
http://www.arches2012.dk

Related event
6th Annual Meeting of the Audiological Research Cores in Europe
26/11/2012 → 27/11/2012
Copenhagen, Denmark
Activity: Attending an event › Participating in or organising a conference

Spontane Otoakustische Emissionen - generiert durch aktive Oszillatoren gruppiert in Frequenzplateaus
Period: 12 Mar 2012
Bastian Epp (Invited speaker)
Department of Electrical Engineering
Hearing Systems

Description
Medautoren:

van Dijk, Pim, Wit, Hero

Related event
Jahrestagung der Deutschen Gesellschaft für Akustik
10/03/2014 → 13/03/2014
Oldenburg, Germany
Activity: Talks and presentations › Conference presentations

35th MidWinter Meeting of the Association for Research in Otolaryngology
Period: Feb 2012
Federica Bianchi (Participant)
Department of Electrical Engineering
Hearing Systems

Description
poster presentation
Documents:
FedericaBianchi_PosterARO

Related event
35th MidWinter Meeting of the Association for Research in Otolaryngology
25/02/2012 → 29/02/2012
San Diego, CA, United States
Activity: Attending an event › Participating in or organising a conference
Assistive Technology for People with Vision and Hearing Impairments
Period: 5 Sep 2010
Hans-Heinrich Bothe (Speaker)
Hearing Systems
Centre for Applied Hearing Research
Department of Electrical Engineering

Description
Audio-visual signal processing – analysis and synthesis of speaking heads

Place: Workshop on Assistive Technology for People with Vision and Hearing Impairments (EU-CWST project). Wroslaw, Poland

Related event
Assistive Technology for People with Vision and Hearing Impairments
05/09/2010 → …
Wroslaw, Poland
Activity: Talks and presentations › Conference presentations

Lippenlesen im Dialog mit dem Computer
Period: 9 Aug 2010
Hans-Heinrich Bothe (Speaker)
Hearing Systems
Centre for Applied Hearing Research
Department of Electrical Engineering

Description
Place: Yearly meeting of the German Association for Hard-of-Hearing People (DSB) Hanover, Germany

Related external organisation
Unknown external organisation
Activity: Talks and presentations › Talks and presentations in private or public companies and organisations

ThinkingHead – eine dialogfähige Computeranimation, die für das Erlernen des Absehens eingesetzt werden kann
Period: 7 Feb 2010
Hans-Heinrich Bothe (Speaker)
Hearing Systems
Centre for Applied Hearing Research
Department of Electrical Engineering

Description
Place: IKT-Forum für Menschen mit Behinderungen: Praxis – Forschung – Entwicklung. Linz, Austria

Related event
IKT-Forum für Menschen mit Behinderungen: Praxis – Forschung – Entwicklung
08/07/2010 → 09/07/2010
Linz, Austria
Activity: Talks and presentations › Conference presentations

Models for the dynamics of articulatory lip movements
Period: 5 Jul 2009
Hans-Heinrich Bothe (Speaker)
Department of Electrical Engineering
Hearing Systems
Centre for Applied Hearing Research

**Description**
Place: Assistive Technology for People with Sensory Impairments: A Hands-On Workshop during ICTA 09. Hammamet, Tunisia

**Related external organisation**
Unknown external organisation
Activity: Talks and presentations › Conference presentations

**Auditory neuroprostheses**
Period: 3 May 2009
Hans-Heinrich Bothe (Speaker)

Hearing Systems
Centre for Applied Hearing Research
Department of Electrical Engineering

**Description**
Place: Yearly meeting of the Audiological Society of Denmark. Copenhagen, Denmark
Degree of recognition: National

**Related event**
Yearly meeting of the Audiological Society of Denmark
03/05/2009 → …
Copenhagen, Denmark
Activity: Talks and presentations › Conference presentations

**ISO TC 43/WG 1: Threshold of hearing (External organisation)**
Period: 1980 → …
Torben Poulsen (Participant)

Department of Electrical Engineering
Hearing Systems

**Description**
International Standardization Organization: Technical Committee 43, Acoustics, Working Group 1: Threshold of hearing
Degree of recognition: International

**Related external organisation**
ISO TC 43/WG 1: Threshold of hearing
Activity: Membership › Membership of committees, commissions, boards, councils, associations, organisations, or similar

**ISO TC43/SC 1/WG 17: Methods of measurement of sound attenuation of hearing protectors (External organisation)**
Period: 1975 → …
Torben Poulsen (Participant)

Department of Electrical Engineering
Hearing Systems
Degree of recognition: International

**Related external organisation**
ISO TC43/SC 1/WG 17: Methods of measurement of sound attenuation of hearing protectors
Activity: Membership › Membership of committees, commissions, boards, councils, associations, organisations, or similar
Prizes:

**ISAAR scholarship**
Helia Relano Iborra (Recipient)
Department of Electrical Engineering, Hearing Systems

**Description**
The ISAAR committee offers a limited number of scholarships to young scientists that would like to participate with a scientific contribution at an ISAAR symposium. The scholarship covers the symposium fee for full participation and accommodation. Travel expenses are not covered. The ISAAR scholarships are intended for young scientists (e.g., PhD-students, post-doctoral students, and others) working in Auditory and Audiological Research or related areas.

**Details**
Awarded date: 2017
Prize: Prizes, scholarships, distinctions

**Outstanding Paper by a Young Presenter**
A. Josefine Sørensen (Recipient)
Department of Electrical Engineering, Hearing Systems

**Details**
Awarded date: Nov 2016
Degree of recognition: International
Granting Organisations: Acoustical Society of America
event: 5th Joint Meeting of the Acoustical Society of America and Acoustical Society of Japan
Prize: Prizes, scholarships, distinctions

**Student Travel Award**
Andreu Paredes Gallardo (Recipient)
Department of Electrical Engineering, Hearing Systems

**Description**
Student Travel Award to attend the 2017 Conference on Implantable Auditory Prostheses (CIAP) at Granlibakken Conference center

**Details**
Awarded date: 16 Jul 2017
event: Conference on Implantable Auditory Prosthesis
Prize: Prizes, scholarships, distinctions

**Young Researcher Award 2016**
A. Josefine Sørensen (Recipient), Lærke Cecilie Bjerre (Recipient) & Thea Mathilde Larsen (Recipient)
Department of Electrical Engineering, Hearing Systems

**Details**
Awarded date: 12 May 2016
Degree of recognition: National
Granting Organisations: Danish Sound Innovation Network
event: Danish Sound Day
Prize: Prizes, scholarships, distinctions

Press clippings:

**Nørderne kommer**
Sébastien Santurette
17/05/2014

**Description**
Interview on singers' vibrato for DR P2
Hearing Systems, Department of Electrical Engineering
Media contribution (1)

Nærderne kommer
17/05/2014
DR P2, Radio
Danmarks Radio
http://www.dr.dk/radio/ondemand/p2/noerderne-kommer-64#!/
Sébastien Santurette
Department of Electrical Engineering, Hearing Systems
Press / Media