

Technical University of Denmark



Relations between perceptual measures of temporal processing, auditory-evoked brainstem responses and speech intelligibility in noise

Papakonstantinou, Alexandra; Strelcyk, Olaf; Dau, Torsten

Published in:
Hearing Research

Link to article, DOI:
[10.1016/j.heares.2011.02.005](https://doi.org/10.1016/j.heares.2011.02.005)

Publication date:
2011

[Link back to DTU Orbit](#)

Citation (APA):

Papakonstantinou, A., Strelcyk, O., & Dau, T. (2011). Relations between perceptual measures of temporal processing, auditory-evoked brainstem responses and speech intelligibility in noise. *Hearing Research*, (1-2), 30-37. DOI: 10.1016/j.heares.2011.02.005

DTU Library

Technical Information Center of Denmark

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research Paper

Relations between perceptual measures of temporal processing, auditory-evoked brainstem responses and speech intelligibility in noise

Alexandra Papakonstantinou^{a,b}, Olaf Strelcyk^{c,1}, Torsten Dau^{c,*}^a International Graduate Research Training Group "Neurosensory Science and Systems", Universität of Oldenburg, Germany^b Medizinische Physik, Universität Oldenburg, D-2611 Oldenburg, Germany^c Centre for Applied Hearing Research, Department of Electrical Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Article history:

Received 11 December 2009

Received in revised form

8 February 2011

Accepted 14 February 2011

Available online xxx

ABSTRACT

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.

© 2011 Published by Elsevier B.V.

1. Introduction

One of the most common complaints of people with cochlear hearing loss concerns difficulty with speech communication. People with mild or moderate losses can typically understand speech reasonably well when they are in a quiet room with only one person talking. However, they experience difficulty when more than one person is talking or when background noise or reverberation is present (e.g., Plomp, 1978, 1994; Bronkhorst, 2000; Helfer and Wilber, 1990). People with severe or profound losses often have problems even when listening to a single talker in a quiet room, and they generally have severe problems when background noise is present. Even if provided with hearing aids, many listeners still experience problems with speech communication particularly

in situations with background noise. Despite the enormous technological progress in hearing-aid technology in the past decade, the benefit from hearing aids seems to vary greatly among individual listeners, even among those showing similar audiograms. This indicates the insufficiency of audibility as a predictor of speech intelligibility in noise (e.g., Dreschler and Plomp, 1985).

There has been considerable controversy in the literature about the reasons for the difficulties in understanding speech (see Moore, 1995, for a review). Some studies have suggested that the difficulties arise primarily from reduced audibility: Absolute thresholds are higher than normal such that the audible proportion of the speech spectrum is reduced for hearing-impaired (HI) listeners (Humes et al., 1987; Zurek and Delhorne, 1987; Lee and Humes, 1995). Others argued that the difficulty understanding speech arises, at least partly, from deficits in the ability to discriminate sounds that are well above absolute threshold (Plomp, 1978; Dreschler and Plomp, 1980, 1985; Glasberg and Moore, 1989). Examples of such deficits include reduced frequency selectivity (e.g., Glasberg and Moore, 1986; Tyler et al., 1986) and deficits in temporal processing abilities (e.g., Fitzgibbons and Wightman, 1982; Tyler et al., 1982; Lorenzi et al., 2006; Hopkins et al., 2008; Strelcyk and Dau, 2009).

Abbreviations: (ABR), Auditory brainstem response; (AM), Amplitude modulation; (SRT), Speech reception threshold; (TFS), Temporal fine structure.

* Corresponding author.

E-mail address: tdau@elektro.dtu.dk (T. Dau).

¹ Present address: Starkey Hearing Research Center, 2150 Shattuck Ave., Suite 408, Berkeley, CA 94704, USA.

In most studies, subjects with a wide range of audiometric configurations were used. Typically, the average degree of hearing loss and the slope of the audiogram varied considerably across subjects. However, several issues may arise when considering such heterogeneous subject groups. For example, it can be difficult to decide whether to compare subject performance at equal sensation level (SL) or equal sound pressure level (SPL). Also, in experiments testing speech intelligibility, the proportion of the speech spectrum that exceeds the absolute threshold varies largely from one subject to another. Furthermore, large variability in the audiometric configuration can obscure effects of other auditory functions such as frequency selectivity as well as monaural and binaural temporal processing. Therefore, in the present study, a homogeneous group of listeners was chosen, minimizing confounding effects of audibility.

The study focused on measures of temporal processing and their relation to speech intelligibility. Temporal processing at low frequencies was addressed via behavioural experiments investigating frequency discrimination, binaural masked detection and amplitude modulation (AM) detection. The first two tasks are supposed to crucially depend on temporal-fine-structure (TFS) information, which refers to the temporal fine structure at the output of the cochlear filters. This fine structure evokes phase-locked activity, i.e., synchronized timing of action potentials, in the subsequent stages of neural processing. Deficits in TFS processing have been found in some earlier studies on, e.g., pure-tone frequency discrimination (e.g., Turner and Nelson, 1982; Freyman and Nelson, 1991), low-rate frequency-modulation detection (e.g., Zurek and Formby, 1981; Moore and Skrodzka, 2002; Buss et al., 2004; Strelcyk and Dau, 2009) and binaural masked detection (e.g., Hall et al., 1984; Gabriel et al., 1992). Also, several studies suggested a relation between ageing and deficits in TFS processing as well as speech reception (e.g., Pichora-Fuller and Schneider, 1992; Strouse et al., 1998; Schneider et al., 2002; Ross et al., 2007; Hopkins et al., 2008; Strelcyk and Dau, 2009). In contrast, AM detection is naturally associated with envelope-based processing. The processing and coding of envelope fluctuations, or amplitude modulations, in the auditory system may be related to the ability to understand speech (e.g., Steeneken and Houtgast, 1980). Mainly modulation frequencies below about 16 Hz have been shown to be crucial for speech intelligibility (e.g., Drullman et al., 1994), but also frequencies above 20 Hz can be important, for example, for the identification of consonants (Christiansen and Greenberg, 2005; Christiansen et al., 2007), particularly in the presence of background noise.

In addition, auditory brainstem responses (ABRs) to transient broadband chirps (Dau et al., 2000) were recorded in the present study. ABRs reflect synchronized neural activity from different levels along the auditory brainstem. In general, the higher the stimulus level the more neurons are activated due to spread of excitation in the cochlea. Typically, stimulus onsets or offsets produce effective responses since many neurons are activated by these transients. However, it is not the spectral content of the input stimuli but the amount of synchronization after cochlear processing that appears to dominate the response amplitude (e.g., Dau et al., 2000; Dau, 2003; Elberling et al., 2007; Wegner and Dau, 2002; Fobel and Dau, 2004). The chirp was originally designed to compensate for cochlear travel time differences across frequencies and was found to produce larger wave-V amplitudes than the traditional click stimulus presented at same stimulation levels (e.g., Dau et al., 2000; Elberling et al., 2007). In the present study, ABR threshold was considered, i.e., the lowest stimulus level at which a repeatable response could be identified, and not ABR at suprathreshold levels. It has been shown recently that the amount of neural across-frequency synchronization can be affected in some listeners despite normal hearing thresholds (Kujawa and Liberman, 2009). Also, it is possible that synchronization plays a role

for encoding weak signals in a noise background, for example through spatio-temporal processing (e.g., Carney et al., 2002; Deng and Geisler, 1987). In the present study, it is hypothesized that a reduced amount of synchronized neural activity at cochlear and brainstem level, which could be reflected in raised ABR thresholds, may affect TFS processing. Therefore, the relation between the behavioural measures of temporal processing and the “objective” electrophysiological measure of temporal processing was considered here.

2. Materials and methods

2.1. Listeners

Overall, twelve subjects participated in the current study. The HI group consisted of one female and six males, ranging between 54 and 81 years in age, with a median age of 66 years. They had bilaterally symmetric high-frequency hearing losses above about 2 kHz, with a difference between left and right thresholds of maximally 20 dB at individual frequencies. The sensorineural origin of the hearing losses was confirmed by means of bone-conduction and tympanometric measurements. Fig. 1 shows the audiometric results of the selected HI ears, identified by a code name. The ears were selected such that the audiograms were most similar across listeners. The test ears were the same in all experiments (binaural masked detection involved both ears). In the following, the HI listeners are ordered according to their pure-tone average thresholds (PTAs) at the frequencies 0.5, 1, 2 and 4 kHz. For example, listener HI1 had a PTA of 35 dB hearing level (HL; ISO 389-8, 2004) and HI7 had a PTA of 48.75 dB HL. At low frequencies, all subjects had pure-tone thresholds within 20 dB HL. The subject group can be considered as homogeneous in terms of the audiogram, but this does not necessarily imply homogeneity in terms of suprathreshold and potential retrocochlear deficits.

The control group of the NH listeners consisted of three males and two females, ranging in age from 24 to 29 years. Thus, this group was not age-matched to the group of the HI listeners in the present study. They had pure-tone thresholds of 10 dB HL or less for frequencies between 125 and 3000 Hz, and 20 dB HL or less for frequencies between 4000 and 8000 Hz.

All listeners received an hourly compensation for their participation in the study and all experiments were approved by the Ethics Committee of Copenhagen County.

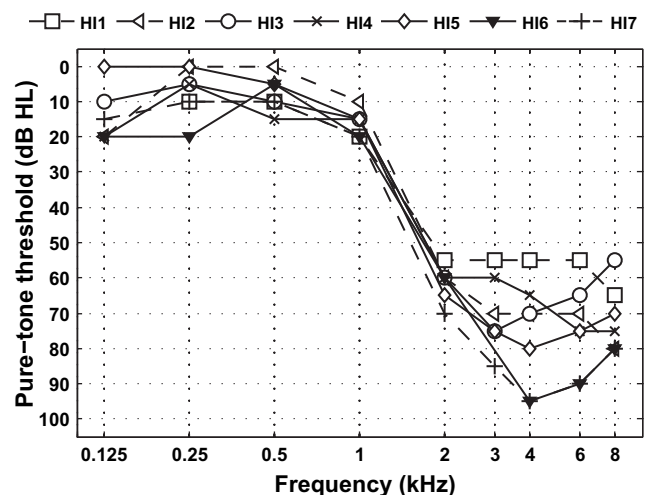


Fig. 1. Pure-tone thresholds for the selected ears of the seven hearing-impaired (HI) listeners with bilaterally symmetrical, high-frequency sensorineural hearing losses.

2.2. Speech intelligibility

SRTs in speech-shaped noise were measured using the DANTALE II speech intelligibility test (Wagener et al., 2003). The SRT was defined as the signal-to-noise ratio at which 50% of the words were correctly identified. The speech was presented at a level of 75 dB SPL while the level of the speech-shaped noise was varied adaptively, starting at a level of 75 dB SPL. Each listener was trained on 60 sentences before data collection. The test was performed in a double-walled, sound attenuating booth. The speech and noise signals were generated in MATLAB and converted to analogue signals using a 16-bit digital-to-analogue (D/A) converter (RME DIGI96/8) at a sampling rate of 44 100 Hz. The signals were presented via Sennheiser HDA200 headphones which were connected to a TDT HB7 audio amplifier.

2.3. Basic psychoacoustical tests

For all of the following tests, a three-interval, three-alternative, forced-choice (3I-3AFC) paradigm was used in combination with a one-up, two-down tracking algorithm converging at the 70.7%-correct point of the psychometric function (Levitt, 1971). Listeners had to indicate the interval which contained the target stimulus and received a visual feedback indicating the correct response. Each run had a total of 12 reversals and thresholds were estimated as the average of the last 8 reversals. For each listener and condition, three runs were taken, and the mean thresholds as well as their standard deviations were calculated from the corresponding three threshold estimates. The results for all basic behavioural tests were obtained after 30 min of training; potential benefits of longer training were not examined. The experiments were carried out using the same equipment as in the speech intelligibility test, with the only difference that Sennheiser HD580 headphones were used here.

Frequency discrimination thresholds were measured for tones at 250 and 1000 Hz, at a level of 75 dB SPL. The tones had a duration of 500 ms and the individual intervals were separated by 250 ms of silence. Listeners had to indicate the interval containing the target tone which had a higher frequency than the reference tone. The initial difference between target and reference frequency was 25%. This difference was either divided by a given step factor following two consecutive correct responses or multiplied by the same factor following a wrong response. The initial step factor was 2 and was multiplied by 0.75 after every second reversal. The minimum step factor was 1.125.

In the *binaural masked detection* experiment, masked thresholds were measured for the diotic N_0S_0 condition with noise and signal presented diotically, as well as for the dichotic N_0S_π condition with diotic noise masker and signal in antiphase. Two signal frequencies were considered: 250 and 1000 Hz. The 480-ms long signal was temporally centred in the 500-ms long noise. The inter-stimulus interval was 250 ms. The masker was two octaves wide, geometrically centred at the signal frequency and had a fixed level of

65 dB SPL. The signal level was varied adaptively, with a final step size of 1 dB.

The *modulation detection* experiment was performed with a sinusoidal carrier at 65 dB SPL. The signal was defined as follows: $s(t) = \sin(2\pi f_c)[1 + m\cos(2\pi f_m t)]$, where f_c indicates the carrier frequency (500 Hz), m denotes the modulation depth (between 0 and 1), and f_m represents the modulation rate (either 8 or 32 Hz). Listeners had to indicate the interval which contained the modulated tone, while the modulation depth was varied adaptively in terms of $20 \times \log(m)$, with a final step size of 1 dB.

2.4. Auditory brainstem responses

ABRs elicited by chirp stimuli were recorded. The chirp developed by Dau et al. (2000) was used which was designed to compensate for cochlear travel-time differences across frequency. The waveform of the chirp was calculated on the basis of the cochlea model by de Boer (1980). The waveforms and (acoustic) magnitude spectra of the chirp are shown in Fig. 2. The chirp duration was 10.34 ms (starting and ending with zero crossings). The chirp had a flat magnitude spectrum corresponding to that of a click stimulus. As a consequence, the chirp started with very small amplitudes at the low-frequency end and increased nonlinearly in amplitude with increasing frequency.

The stimuli for the NH and HI listeners were played back through EAR TONE 3A insert earphones. The presentation of the chirp stimuli was temporally jittered, with a mean presentation rate of 22 Hz (the resulting inter-stimulus intervals were equally distributed between 37 and 59 ms). The stimuli were presented at five different levels: 18, 24, 30, 36 and 42 dB SL. Here, the behavioural thresholds of the individual listeners, represented as peak equivalent sound pressure levels (pSPL), were determined in a 3I-3AFC experiment. Each ABR recording session started with the highest stimulation level of 42 dB SL. The level was then lowered successively in 6-dB steps.

Listeners lay on a couch in an acoustically and electrically shielded room. The ABRs were measured differentially between electrodes applied to the vertex (Cz in the 10/20 system) and the ipsilateral mastoid (M1 or M2). Another electrode, applied to the forehead (Fpz), served as ground. The electrode signals were acquired using a Nicolet spirit electrodiagnostic system. They were processed by a 2nd-order Bessel low-pass digital filtering with a cut-off frequency of 3 kHz. For each stimulus condition, two recordings were obtained, each consisting of 2000 sweeps.

Evoked potential estimates were obtained by averaging the 2000 sweeps and wave Vs were detected by visual inspection of these evoked potential estimates. Wave-V peak-to-peak amplitude was measured from the peak to the largest negativity following it. The ABR wave-V threshold was defined as the lowest stimulus level at which a repeatable wave V could be identified in the response waveform. In addition, the following criteria had to be met: (i) wave-V peak had to fall within the time window of 3–8 ms after

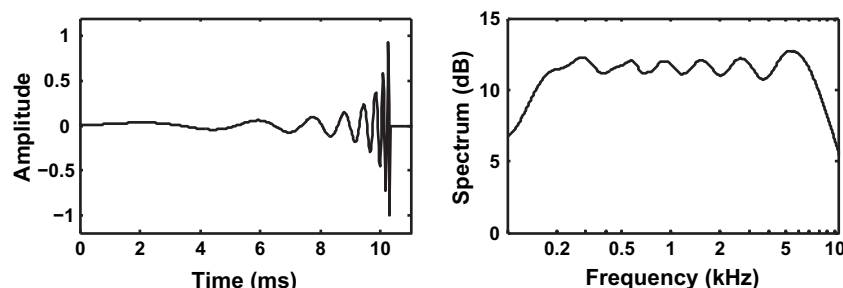


Fig. 2. Left panel: Waveform of the chirp stimulus used in the ABR experiments, with a duration of 10.34 ms. Right panel: Acoustic magnitude spectrum of the chirp.

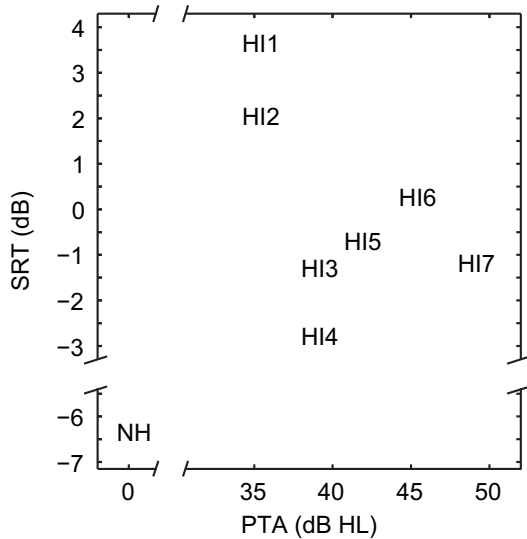


Fig. 3. Speech reception threshold (SRT) in 75 dB noise as a function of the PTA (average pure-tone threshold among 0.5, 1, 2 and 4 kHz). Results for the HI listeners are represented by the corresponding code names. “NH” represents the single native Danish speaker from the NH group. The abscissa and ordinate are broken for better visualization of the data.

chirp offset; (ii) wave-V peak had to be reproducible for the two independent recordings, for each stimulus condition, and (iii) the shift in wave-V latency between the threshold level and the higher levels had to be smaller than 3 ms.

3. Results

3.1. Speech intelligibility

Fig. 3 shows the SRT results as a function of the PTA for the HI and NH listeners. As can be seen, SRTs varied strongly among the HI listeners, with values between -3 and $+4$ dB signal-to-noise ratio. The NH reference value shows an SRT of -6.5 dB. In this particular experiment, only one NH listener was considered because this person was the only native Danish speaker in this group. The SRT obtained by this listener is slightly higher than the average value of -8.4 dB obtained in Wagener et al. (2003). The group of the HI

listeners of the present study was considered to be homogeneous in terms of the audiograms; nevertheless, some variability in the amount of hearing losses remained, mainly at high frequencies. However, SRTs (for the HI listeners) were not significantly correlated with the PTAs (Pearson correlation and two-tailed p -value: $R = -0.5$, $p > 0.05$). Hence, the variability within the audiograms cannot account for the observed variability in SRTs. In fact, listener HI1 had large difficulties understanding speech in noise ($+4$ dB signal-to-noise ratio) even though this subject had the highest pure-tone sensitivity and had reasonable access to the high-frequency parts of the speech signal since the signal level was above threshold. In contrast, subject HI7 which had the largest sensitivity loss showed relatively good ability to understand speech, with an SRT at a signal-to-noise ratio of -1.3 dB.

3.2. Frequency discrimination

Fig. 4 shows the frequency discrimination thresholds (FDTs) at 250 Hz (left panel) and 1000 Hz (right panel) for the HI (code names) and the NH listeners (circles). The horizontal black lines denote the mean FDTs and the corresponding boxes represent ± 1 standard deviation (SD) for the NH (white) and HI (grey) listeners, respectively. The FDTs are plotted on a logarithmic frequency scale, with the left and right ordinate indicating the FDTs as proportion of the signal frequency and in Hz, respectively. The results for the NH listeners are similar for the two frequencies which are consistent with previous studies (e.g., Sek and Moore, 1995). The thresholds did not differ significantly between the NH and HI listeners (two-tailed t -test at 250 Hz: $p > 0.05$; at 1 kHz: $p > 0.05$). For the HI listeners, a significant correlation between FDTs and SRTs was observed at both frequencies (250 Hz: $R = 0.96$, $p < 0.001$; 1 kHz: $R = 0.90$, $p < 0.005$). The listener with the worst performance in the speech intelligibility tasks (HI1) also showed the worst performance in the frequency discrimination task. The listener with the best speech performance actually showed FDTs in the range of values obtained in the normal-hearing listeners. FDTs were not significantly correlated with individual hearing thresholds, neither at the single frequencies of 250 Hz and 1 kHz nor in terms of the PTA ($p > 0.05$). Therefore, the correlation between FDTs and SRTs remained significant when individual hearing thresholds were partialled out (250 Hz: $R_{\text{partial}} = 0.96$, $p < 0.005$; 1 kHz: $R_{\text{partial}} = 0.91$, $p < 0.005$).

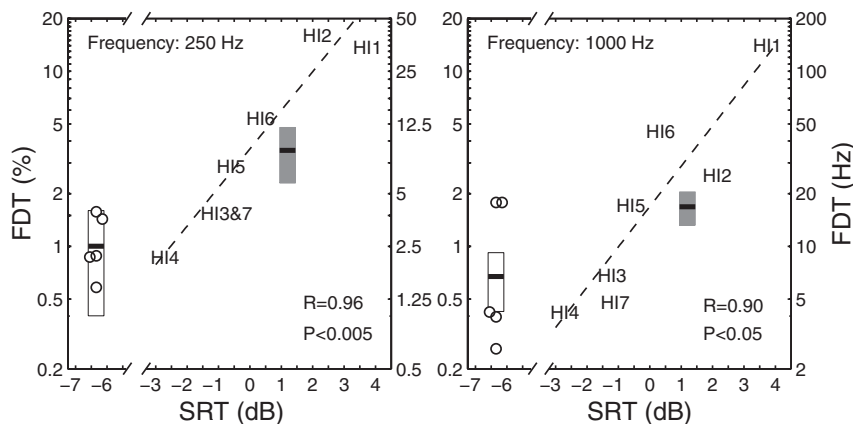


Fig. 4. Frequency discrimination thresholds (FDTs) at 250 Hz (left panel) and 1000 Hz (right panel) signal frequency as a function of speech reception threshold (SRT) in 75 dB noise. The results for the HI listeners are indicated by the corresponding code names while the results for the NH listeners are represented as circles. The horizontal black lines denote the mean FDT and the corresponding boxes represent the ± 1 standard deviation (SD) for the NH (white) and HI (grey) listeners. The left abscissa expresses FDT as a proportion of the signal frequency and the right abscissa indicates FDT in Hz; both are shown on a logarithmic scale. The dashed lines indicate the robust regression line between FDT and SRT. The abscissa is interrupted to permit better visualization of the data.

3.3. Binaural masked detection

Fig. 5 shows the binaural masked thresholds for the dichotic N_0S_π condition at 250 (left panel) and 1000 Hz (right panel) as a function of the SRT. The results for the HI and NH listeners are shown in a similar way as in Fig. 4. For 1 kHz (right panel), the HI listeners showed significantly higher N_0S_π thresholds than the NH listeners (two-tailed t -test: $p = 0.04$) whereas the N_0S_0 thresholds did not differ significantly between the two groups ($p > 0.05$; not shown). Nevertheless, the binaural masking level differences, i.e., the differences between N_0S_π and N_0S_0 thresholds, did not differ significantly between the two groups ($p > 0.05$). The N_0S_π thresholds at 1000 Hz were significantly correlated with SRTs ($R = 0.76$, $p < 0.05$). This correlation was no longer significant when hearing thresholds at 1 kHz were partialled out ($p = 0.06$). For 250 Hz (left panel), there was no significant difference between the N_0S_π thresholds obtained for the NH and HI listeners. Also, at this frequency, the N_0S_π thresholds were only marginally correlated with SRT ($R = 0.70$, $p = 0.08$). Here, it should be kept in mind that the SRTs were measured only for one ear. This may explain why the correlation found here between dichotic masked detection and monaural speech reception was weaker than the correlation between monaural frequency discrimination and speech reception discussed above. Finally, a correlation between the N_0S_π thresholds and the monaural FDTs at 1 kHz just reached significance ($R = 0.75$, $p = 0.05$).

3.4. Amplitude modulation detection with a sinusoidal carrier

While the frequency discrimination and binaural masked detection tasks are commonly associated with the ability to process TFS information, this experiment dealt with the perception of the envelope fluctuations of a sound. Fig. 6 shows the AM detection thresholds for 8 Hz (left panel) and 32 Hz (right panel), imposed on a 500-Hz carrier, as a function of the SRT. The results for the HI and NH listeners are indicated in a similar way as in Figs. 4 and 5. AM detection thresholds were not significantly different for the NH and HI listeners at both modulation frequencies ($p > 0.05$). In fact, most of the HI listeners showed thresholds close to normal at both rates of 8 Hz and 32 Hz, with the highest threshold at about -15 dB. This is consistent with results from earlier studies on amplitude modulation detection in sensorineural HI listeners (e.g., Bacon and Gleitman, 1985). Furthermore, the AM detection thresholds were not significantly correlated with the SRTs (8 Hz: $R = 0.21$, $p > 0.05$; 32 Hz: $R = 0.48$, $p > 0.05$). Also, the AM detection thresholds were not correlated with the individual hearing threshold at 500 Hz ($p > 0.05$).

3.5. Auditory brainstem responses

The upper left panel of Fig. 7 shows the chirp-evoked ABR wave-V thresholds (in peSPL) for the individual listeners, including the normal-hearing listeners, as a function of the SRT. The two black horizontal bars and the corresponding white and grey boxes around them indicate the mean and ± 1 standard deviations for the normal and the hearing-impaired listeners, respectively. Unfortunately, no reliable response could be obtained for listener HI7, i.e., the recordings were very noisy and a large number of recorded epochs was rejected. Thus, since responses from only six hearing-impaired listeners could be obtained in this study, it was not possible to perform a statistical correlation analysis between the ABR results and the results from the other measures. Nevertheless, the observed trends indicating possible relations between the different outcome measures will be discussed in the following. Thus, increasing wave-V thresholds tended to be associated with increasing SRTs, i.e., decreasing speech intelligibility. Wave Vs could not be detected for listener HI2. Therefore, for this subject, wave-V threshold was represented by the maximum stimulation level of 120 dB peSPL in the figure.

The upper right panel of Fig. 7 shows wave-V threshold as a function of the N_0S_π thresholds at 1 kHz for the individual listeners. It can be seen that larger ABR thresholds tend to be associated with larger binaural thresholds. The lower left panel shows the ABR threshold as a function of the FDT that was averaged between 250 Hz and 1 kHz for each individual listener. Increasing ABR thresholds seem to be associated with increasing FDTs. Finally, the lower right panel shows the ABR threshold as a function of the AM detection threshold for 8 Hz. There is no trend indicating a relation between evoked potential amplitude and modulation detection sensitivity. A similar result was observed for the modulation frequency of 32 Hz (not shown in the figure).

4. Discussion

The hypothesis of the current study was that speech reception in noise partly depends on the processing of temporal information at low frequencies. It was examined whether low-frequency temporal processing (≤ 1 kHz) could be affected in regions of normal hearing, for listeners with (steeply sloping) high-frequency hearing losses at higher frequencies, and whether potential deficits could be related to speech reception performance in noise. Furthermore, it was investigated, whether objective ABR wave-V thresholds obtained for the chirp stimuli were correlated with the behavioural measures of temporal processing.

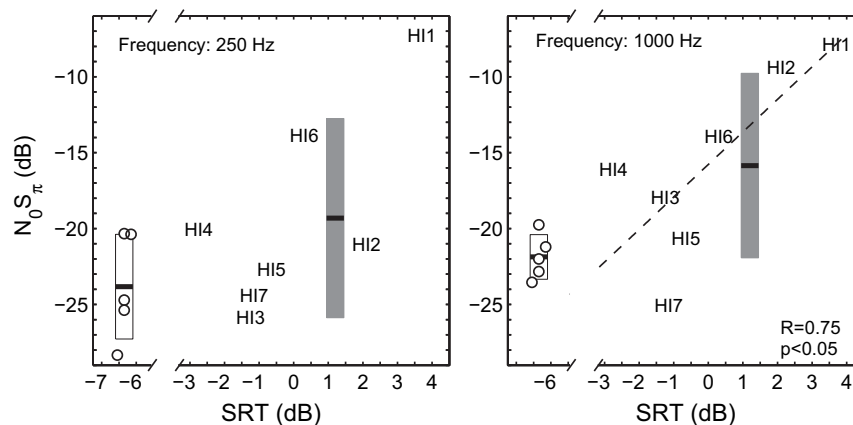


Fig. 5. Binaural masked detection threshold, N_0S_π , for signal frequencies 250 Hz (left panel) and 1000 Hz (right panel) as a function of speech reception threshold (SRT). The dashed line indicates the robust regression line between N_0S_π threshold and SRT at 1000 Hz. The abscissa is broken for better visualization of the data.

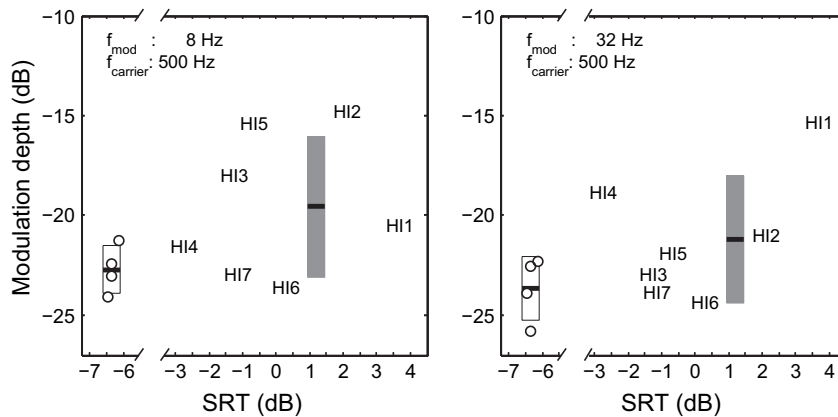


Fig. 6. Modulation detection thresholds for a sinusoidal carrier at 500 Hz for modulation frequencies of 8 (left panel) and 32 Hz (right panel) are shown as a function of SRT. The abscissa is broken for better visualization of the data.

Most of the HI listeners showed poorer results than the NH listeners both in the behavioural experiments testing frequency discrimination, binaural masked detection, and speech reception as well as in the objective measure using ABRs wave-V amplitude as an indicator of neural (temporal) synchronization across frequency. Fig. 8 illustrates the relationships observed in the present study. Hearing thresholds were not associated with any of the other measures. To some extent, this can be attributed to the fact that the HI listeners had similar audiograms and essentially normal (or close to normal) sensitivity at low frequencies. Regarding the speech results, high-frequency consonant information may have been inaudible to *all* of the HI listeners. This would explain why the measured SRTs did not reflect the across-subject variability within the high-frequency pure-tone thresholds. For the HI listeners, a correlation was observed

between monaural frequency discrimination and binaural masked detection at 1 kHz, both supposed to rely on the ability to process TFS. These correlations are indicated in Fig. 8 as connecting lines between the measures (no arrowheads are shown since the order of processing in the auditory system is unclear). Furthermore, frequency discrimination was correlated with speech reception in noise, consistent with previous studies (Tyler et al., 1983; Glasberg and Moore, 1989; Noordhoek et al., 2001). Also, binaural masked N_0S_{π} thresholds were correlated with the SRTs.

Regarding the ABR results of the present study, the chirp stimulation (at corresponding sensation levels) was less effective for the HI listeners than for the NH listeners. One reason for this could be due to alterations of cochlear travel times in the HI listeners as proposed in earlier ABR studies (Don et al., 1998; Strelcyk et al.,

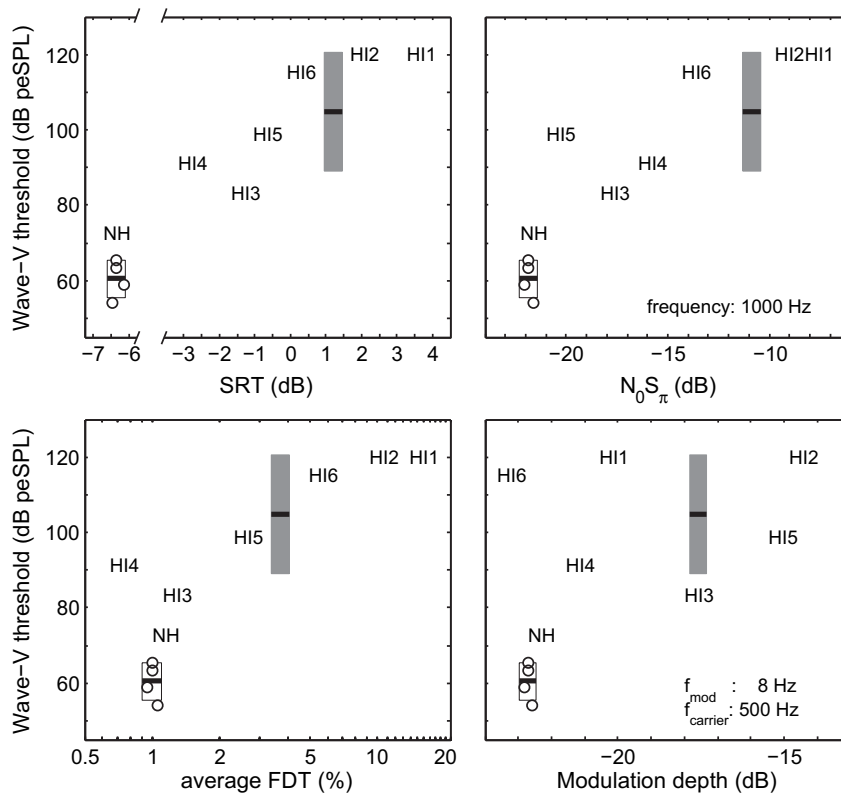


Fig. 7. Chirp-evoked ABR wave-V threshold as a function of SRT (upper left panel), binaural masked detection threshold, N_0S_{π} , at 1 kHz signal frequency (upper right panel), frequency discrimination threshold, FDT, averaged across 250 Hz and 1 kHz (lower left panel) and modulation detection threshold for a 500-Hz carrier modulated at 8 Hz (lower right panel). Wave V was not measurable for HI7 listener due to a large high-number of rejected epochs.

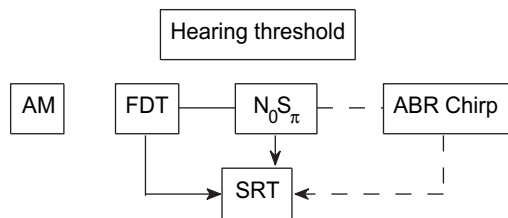


Fig. 8. Relations between the results for the different tests for the group of HI listeners: Pure-tone hearing thresholds (hearing threshold), amplitude modulation detection (AM), frequency discrimination (FDT), dichotic binaural masked detection (N_0S_π), speech reception (SRT) and chirp-evoked auditory brainstem responses (ABR chirp). Significant correlations between measures are indicated as solid connecting lines. Dashed lines indicate a relation between two measures; since ABR data could only be obtained from six listeners, no correlation analysis could be performed here. When no arrowheads are shown, this indicates that the order of processing in the auditory system cannot be specified.

2009). Such altered travel times would typically be associated with changes in frequency selectivity as one of the potential consequences of a cochlear hearing impairment. Alternatively, or in addition, HI listeners might not profit from chirp stimulation in the same way as NH listeners since the neural activity across frequency could be less synchronized in the HI listeners even in the case of “normal” cochlear delays. Interestingly, a trend was observed for ABR wave-V thresholds to increase with increasing SRTs, i.e., decreasing speech reception performance. Hence, the present ABR results may suggest that temporal synchronization of neural activity (at low-frequencies) at the level(s) of processing responsible for wave-V generation, can be affected in some of the HI listeners and that this, in turn, can affect speech intelligibility in noise. This hypothesis is further supported by the finding that chirp-evoked wave-V thresholds also tended to be associated with FDTs (at 250 Hz and 1 kHz) and N_0S_π thresholds (at 1 kHz), both representing measures of TFS processing. These trends are indicated by the dashed lines in Fig. 8. However, in order to allow for a more solid statistical analysis of the relations between the ABR results and the other outcome measures, data from a larger group of listeners would be required.

Several studies suggested a relation between ageing and deficits in TFS processing as well as speech reception (e.g., Pichora-Fuller and Schneider, 1992; Strouse et al., 1998; Schneider et al., 2002; Ross et al., 2007; Hopkins et al., 2008; Strelcyk and Dau, 2009). It has also been demonstrated that human spiral ganglion cells decline with age (Otte et al., 1978) and this can be seen in areas remote from regions of threshold elevation (Felder and Schrott-Fischer, 1995). However, in the present study, age did not correlate with the results from the different outcome measures ($p > 0.05$). Also, a listener’s age *per se* may not be particularly informative since it does not really clarify or explain relations between outcome measures as investigated in the present study.

The NH and HI listeners showed *similar* performance on the AM detection task, indicating that the ability to process envelope information at low (carrier) frequencies might *not* be affected by the high-frequency hearing losses of the listeners. Also, the ability to detect AM was neither related to speech reception in noise nor to any of the other measures in this study. The former is consistent with findings from previous studies that showed little or no correlation between envelope-based temporal resolution measures and speech intelligibility (e.g., Festen and Plomp, 1983; Dubno and Dirks, 1990; van Rooij and Plomp, 1990; Takahashi and Bacon, 1992). For deterministic carriers (such as the pure tones used here), changes in the amount of cochlear compression, commonly associated with a sensorineural hearing impairment, do not affect or sometimes even improve amplitude modulation detection, when the comparison with NH listeners is made at comparable

sensation levels (e.g., Moore, 1995). Here, the carrier frequency (500 Hz) fell in the range of normal hearing.

The performance of some of the HI listeners in the present study bears similarity to what has been observed in people with auditory neuropathy (e.g., Rance et al., 2004; Zeng et al., 2005), albeit with less extreme performance abnormalities: The listeners HI1 and HI2 showed reduced ABR responses (as indicated by elevated ABR thresholds) and reduced performances in speech perception, frequency discrimination, AM detection, and fine-structure-based binaural processing. Hence, the auditory processing deficits observed in these two listeners may have had similar causes as those observed in listeners diagnosed with auditory neuropathy.

Overall, the results from the present study suggest that deficits in the processing of low-frequency TFS information may affect speech reception in noise. Furthermore, they may suggest a relationship between TFS processing and ABR thresholds. However, additional data from a larger group of listeners are needed to confirm the significance of these relations. Eventually, a better understanding of the proposed relationships might help to provide persons with high-frequency hearing losses with better diagnosis and treatment options.

Acknowledgements

This study was supported by GN ReSound, Oticon and Widex. We thank Arne Nørby Rasmussen of the Rigshospitalet Copenhagen for his technical assistance in the measurements and his contribution in providing us with suitable test listeners. We also thank the associate editor Fan-Gang Zeng and the two anonymous reviewers for their constructive and helpful comments and suggestions.

References

- Bacon, S.P., Gleitman, R.M., 1985. Modulation detection in subjects with relatively flat hearing losses. *J. Speech Hear. Res.* 35, 642–653.
- Bronkhorst, A.-W., 2000. The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. *Acustica* 86, 117–128.
- Buss, E., Hall, J.-W., Grose, J.-H., 2004. Temporal fine-structure cue to speech and pure tone modulation in observers with sensorineural hearing loss. *Ear Hear.* 25, 242–250.
- Carney, L.-H., Heinz, M.-G., Evilsizer, M.-E., Gilkey, R.-H., Colburn, H.-S., 2002. Auditory phase opponency: A temporal model for masked detection at low frequencies. *Acta Acust. United Ac.* 88, 334–347.
- Christiansen, T.-U., Greenberg, S., 2005. Frequency selective filtering of the modulation spectrum and its impact on consonant identification. In: *Twenty First Danavox Symposium*, pp. 585–599.
- Christiansen, T.-U., Dau, T., Greenberg, S., 2007. Spectro-temporal processing of speech – an information-theoretic framework. In: Kollmeier, B., Klump, G., Hohmann, V., Langemann, U., Mauermann, M., Uppenkamp, S., Verhey, J. (Eds.), *Hearing – From Sensory Processing to Perception*. Springer Verlag, Berlin, pp. 515–523.
- Dau, T., Wegner, O., Mellert, V., Kollmeier, B., 2000. Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion. *J. Acoust. Soc. Am.* 107, 1530–1540.
- Dau, T., 2003. The importance of cochlear processing for the formation of auditory brainstem and frequency following responses. *J. Acoust. Soc. Am.* 113, 936–950.
- de Boer, E., 1980. Auditory physics. *Physical principles in hearing theory I*. Phys. Rep. 62, 87–174.
- Deng, L., Geisler, C.-D., 1987. A composite auditory model for processing speech sounds. *J. Acoust. Soc. Am.* 82, 2001–2012.
- Don, M., Ponton, C.-W., Eggermont, J.-J., Kwong, B., 1998. The effects of sensory hearing loss on cochlear filter times estimated from auditory brainstem response latencies. *J. Acoust. Soc. Am.* 104, 2280–2289.
- Dreschler, W.-A., Plomp, R., 1980. Relation between psychophysical data and speech perception for hearing-impaired subjects. I. *J. Acoust. Soc. Am.* 68, 1608–1615.
- Dreschler, W.-A., Plomp, R., 1985. Relations between psychophysical data and speech perception for hearing-impaired subjects. II. *J. Acoust. Soc. Am.* 78, 1261–1270.
- Drullman, R., Festen, J.-M., Plomp, R., 1994. Effect of reducing slow temporal modulations on speech reception. *J. Acoust. Soc. Am.* 95, 2670–2680.
- Dubno, J.-R., Dirks, D.-D., 1990. Associations among frequency and temporal resolution and consonant recognition for hearing-impaired listeners. *Acta Otolaryngol. Suppl.* 469, 23–29.

- Elberling, C., Don, M., Cebulla, M., Stürzebecher, E., 2007. Auditory steady-state responses to chirp stimuli based on cochlear traveling wave delay. *J. Acoust. Soc. Am.* 122, 2772–2785.
- Felder, E., Schrott-Fischer, A., 1995. Quantitative evaluation of myelinated nerve fibres and hair cells in cochleae of humans with age-related high-tone hearing loss. *Hear. Res.* 91, 19–32.
- Festen, J.-M., Plomp, R., 1983. Relation between auditory functions in impaired hearing. *J. Acoust. Soc. Am.* 73, 652–662.
- Fitzgibbons, P.-J., Wightman, F.-L., 1982. Gap detection in normal and hearing-impaired listeners. *J. Acoust. Soc. Am.* 72, 761–765.
- Fobel, O., Dau, T., 2004. Searching for the optimal stimulus eliciting auditory brainstem in humans. *J. Acoust. Soc. Am.* 116, 2213–2222.
- Freyman, R.-L., Nelson, D.-A., 1991. Frequency discrimination as a function of signal frequency and level in normal-hearing and hearing-impaired listeners. *J. Speech Hear. Res.* 34, 1371–1386.
- Gabriel, K.-J., Koehnke, J., Colburn, H.-S., 1992. Frequency dependence of binaural performance in listeners with impaired binaural hearing. *J. Acoust. Soc. Am.* 91, 336–347.
- Glasberg, B.-R., Moore, B.-C.-J., 1986. Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *J. Acoust. Soc. Am.* 79, 1020–1033.
- Glasberg, B.-R., Moore, B.-C.-J., 1989. Psychoacoustic abilities of subjects with unilateral and bilateral cochlear impairment and their relationship to the ability to understand speech. *Scand. Audiol.* 32, 1–25.
- Hall, J.-W., Tyler, R.-S., Fernandes, M.-A., 1984. Factors influencing the masking level difference in cochlear hearing-impaired and normal-hearing listeners. *J. Speech Hear. Res.* 27, 145–154.
- Helfer, K.S., Wilber, L.A., 1990. Hearing loss, aging, and speech perception in reverberation and noise. *J. Speech Hear. Res.* 33, 149–155.
- Hopkins, K., Moore, B.-C.-J., Stone, M.-A., 2008. Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *J. Acoust. Soc. Am.* 123, 1140–1153.
- Humes, L.-E., Dirks, D.-D., Bell, T.-S., Kincaid, G.-E., 1987. Recognition of nonsense syllables by hearing-impaired listeners and by noise-masked normal hearers. *J. Acoust. Soc. Am.* 81, 765–773.
- ISO 389-8, 2004. Acoustics—Reference Zero for the Calibration of Audiometric Equipment—Part 8: Reference Equivalent Threshold Sound Pressure Levels for Pure Tones and Circumaural Earphones. International Organization for Standardization, Geneva.
- Kujawa, S.-G., Liberman, M.-C., 2009. Adding insult to injury: Cochlear nerve degeneration after 'Temporary' noise-induced hearing loss. *J. Neurosci.* 29, 14077–14085.
- Lee, L.-W., Humes, L.-E., 1995. Loudness functions in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 97, 3345.
- Levitt, H., 1971. Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–477.
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., Moore, B.-C.-J., 2006. Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc. Natl. Acad. Sci. U. S. A.* 103, 18866–18869.
- Moore, B.-C.-J., 1995. *Perceptual Consequences of Cochlear Damage*. Oxford University Press, Oxford.
- Moore, B.-C.-J., Skrodzka, E., 2002. Detection of frequency modulation by hearing-impaired listeners: effects of carrier frequency, modulation rate, and added amplitude modulation. *J. Acoust. Soc. Am.* 111, 327–335.
- Noordhoek, I.-M., Hougast, T., Festen, J.-M., 2001. Relations between intelligibility of narrow-band speech and auditory functions, both in the 1-kHz frequency region. *J. Acoust. Soc. Am.* 109, 1197–1212.
- Otte, J., Schuknecht, H.F., Kerr, A.G., 1978. Ganglion cell populations in normal and pathological human cochleae. Implications for cochlear implantation. *Laryngoscope LXXXVIII*, 1231–1246.
- Pichora-Fuller, K.M., Schneider, B.A., 1992. The effect of interaural delay of the masker on masking-level differences in young and old adults. *J. Acoust. Soc. Am.* 91, 2129–2135.
- Plomp, R., 1978. Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J. Acoust. Soc. Am.* 63, 533–549.
- Plomp, R., 1994. Noise, amplification, and compression: considerations of three main issues in hearing aid design. *Ear Hear.* 15, 2–12.
- Rance, G., McKay, C., Grayden, D., 2004. Perceptual characterization of children with auditory neuropathy. *Ear Hear.* 25, 34–46.
- Ross, B., Fujioka, T., Tremblay, K.L., 2007. Aging in binaural hearing begins in mid-life: evidence from cortical auditory-evoked responses to changes in interaural phase. *J. Neurosci.* 27, 11172–11178.
- Schneider, B.A., Daneman, M., Pichora-Fuller, M.K., 2002. Listening in aging adults: from discourse comprehension to psychoacoustics. *Can. J. Exp. Psychol.* 56, 139–152.
- Sek, A., Moore, B.-C.-J., 1995. Frequency discrimination as a function of frequency, measured in several ways. *J. Acoust. Soc. Am.* 97, 2479–2486.
- Steeneken, H.-J.-M., Houtgast, T., 1980. A physical method for measuring speech-transmission quality. *J. Acoust. Soc. Am.* 67, 318–326.
- Strelcyk, O., Dau, T., 2009. Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *J. Acoust. Soc. Am.* 125, 3328–3345.
- Strelcyk, O., Christoforidis, D., Dau, T., 2009. Relation between derived-band auditory brainstem response latencies and behavioral frequency selectivity. *J. Acoust. Soc. Am.* 126, 1878–1888.
- Strouse, A., Ashmed, D.H., Ohde, R.N., Grantham, D.W., 1998. Temporal processing in the aging auditory system. *J. Acoust. Soc. Am.* 104, 2385–2399.
- Takahashi, G.A., Bacon, S.P., 1992. Modulation detection, modulation masking, and speech understanding in noise in the elderly. *J. Speech Hear. Res.* 35, 1410–1421.
- Turner, C.W., Nelson, D.A., 1982. Frequency discrimination in regions of normal and impaired sensitivity. *J. Speech Hear. Res.* 25, 34–41.
- Tyler, R.S., Summerfield, Q., Wood, E.J., Fernandes, M.A., 1982. Psychoacoustic and phonetic temporal processing in normal and hearing-impaired listeners. *J. Acoust. Soc. Am.* 72, 740–752.
- Tyler, R.S., Wood, E.J., Fernandes, M., 1983. Frequency resolution and discrimination of constant and dynamic tones in normal and hearing-impaired listeners. *J. Acoust. Soc. Am.* 74, 1190–1199.
- Tyler, R.S., Preece, J.P., Gantz, B.J., Otto, S.R., 1986. Previous experience as a confounding factor in comparing cochlear-implant processing schemes. *J. Speech Hear. Res.* 29, 282–287.
- van Rooij, J.C.G.M., Plomp, R., 1990. Auditive and cognitive factors in speech perception by elderly listeners. II: multivariate analyses. *J. Acoust. Soc. Am.* 88, 2611–2624.
- Wagener, K., Josvassen, J.L., Ardenkjær, R., 2003. Design, optimization and evaluation of a Danish sentence test in noise. *Int. J. Audiol.* 42, 10–17.
- Wegner, O., Dau, T., 2002. Frequency specificity of chirp-evoked auditory brainstem responses. *J. Acoust. Soc. Am.* 111, 1318–1329.
- Zeng, F.G., Kong, Y.Y., Michalewski, H.J., Starr, A., 2005. Perceptual consequences of disrupted auditory nerve activity. *J. Neurophysiol.* 93, 3050–3063.
- Zurek, P.M., Formby, C., 1981. Frequency-discrimination ability of hearing-impaired listeners. *J. Speech Hear. Res.* 24, 108–112.
- Zurek, P.-M., Delhorne, L.-A., 1987. Consonant reception in noise by listeners with mild and moderate sensorineural hearing impairment. *J. Acoust. Soc. Am.* 82, 1548–1559.