

Impaired auditory functions underlying degraded speech perception in noise

OLAF STRELCYK AND TORSTEN DAU

Centre for Applied Hearing Research, Ørsted•DTU, Technical University of Denmark, DK-2800 Lyngby, Denmark

Hearing-impaired people often experience great difficulty with speech communication when background noise is present. In most cases, the problem persists even if reduced audibility has been compensated for by hearing aids. Clearly, other impairment factors besides reduced audibility 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 speech intelligibility in the presence of stationary as well as fluctuating interferers, the individual's frequency selectivity and the integrity of temporal fine-structure processing. The latter was addressed by measuring the lateralization threshold for low-frequency tones with ongoing interaural phase delays. In addition, this lateralization threshold was measured in a stationary noise background in order to assess the persistence of the fine-structure processing to interfering noise. This may play a crucial role for the ability to listen into the dips of fluctuating background interferers.

INTRODUCTION

A sensorineural hearing impairment manifests itself not only in an elevation of absolute hearing thresholds but also in changes of sound perception well above threshold (Moore, 1996). The present study investigated performance on selected psychoacoustical supra-threshold tasks in a common group of hearing-impaired listeners. This was done in order to test the integrity of basic monaural and binaural auditory functions in the individual per se and to throw light on possible relations between these functions and speech perception performance. Groups of listeners with homogeneous audiograms were selected to minimize confounding effects otherwise introduced by varying audibility. In this way, it should also be possible to study mutual relations between the basic auditory functions and to draw reliable conclusions based on a relatively small number of subjects. Performance was measured in speech reception, frequency selectivity, binaural masked detection and binaural lateralization tasks. Frequency selectivity was included because relations between frequency selectivity and speech perception, particularly in noise, have been reported previously (e.g., Dreschler and Plomp, 1985; Horst, 1987). Recently, the processing of temporal fine-structure in hearing-impaired listeners has received considerable attention. Deficits in fine-structure coding might be partly responsible for problems with understanding speech, particularly in noise (Buss *et al.*, 2004; Lorenzi *et al.*, 2006). Therefore, two binaural tests were included here as sensitive indicators of deficits in fine-structure processing: masking level differences and tone lateralization thresholds based on ongoing phase delays. In

addition, the latter was also assessed in the presence of background noise interferers in order to test the persistence of the phase-locking dependent fine-structure cue. This might be crucial in a real listening situation where a certain level of noise background will always be present. Apart from the speech perception measure, all tasks were performed at 750 Hz, where audibility was essentially normal for all subjects participating in the study. Since low-frequency processing is important for binaural as well as monaural tasks (such as for example the detection of mistuned harmonics), we were particularly interested in a possible impact of hearing impairment on the auditory processing at low frequencies, despite normal audiometric thresholds.

METHODS

This section describes the methods used in order to measure speech reception, frequency selectivity, binaural masking level differences and lateralization thresholds.

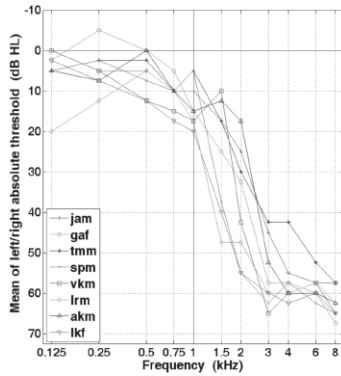


Fig. 1: Mean of left and right audiograms for the eight HI listeners

Listeners

The 16 listeners who participated in this study had bilaterally symmetric (within 10 dB, exceptions see below) audiograms. They are categorized into the following three groups. (i) Six normally hearing (NH) listeners, aged between 20 to 55 years (median: 24), served as control group, having audiometric thresholds within 15 dB HL (re ISO 389-8) at all the tested frequencies from 125 Hz to 8 kHz. (ii) Eight listeners with moderately impaired hearing (HI) were selected to form a homogeneous group in terms of their audiograms (steeply sloping between 1 and 3 kHz), which are shown in Fig. 1. They ranged in age from 24 to 74 (median: 63). The sensorineural origin of their hearing loss was established by means of bone-conduction measurements, tympanometry and otoscopy. (iii) The remaining two subjects included in this study showed normal audiometric thresholds (within 10 dB HL at all frequencies). However, both of them, *lvf* (46) and *kwf* (26) complained about difficulties with understanding speech in noisy backgrounds. Therefore, we classified them as having an ‘obscure dysfunction’ (OD). Additionally, all subjects were screened on a binaural pitch task, testing the ability to hear a Huggins’ pitch C-scale (see Santurette and Dau, 2007). All of them perceived the pitch, thereby suggesting that a severe central auditory deficiency – as conjectured

by Santurette and Dau – was absent.

Stimuli

Speech reception

Speech reception thresholds (SRT) were measured with DANTALE II, a Danish closed-set sentence test (Wagener *et al.*, 2003), in the presence of three different kinds of background noise: stationary speech-shaped noise, sinusoidally amplitude modulated speech-shaped noise (fully modulated at 8 Hz) and dichotic, lateralized noise. The latter was generated by introducing an interaural time delay (ITD) of 740 μ s to the speech-shaped noise. Here, we denote the gain relative to the SRT for speech-shaped noise as spatial masking release.

Frequency selectivity

Auditory filter shapes at 750 Hz were determined for both ears using a notched-noise paradigm (Patterson and Nimmo-Smith, 1980). The 750-Hz target tones of 440 ms duration were presented at a fixed level of 50 dB SPL, and were temporally centered in the 550 ms noise maskers. Maskers and tones were \cos^2 gated with ramp durations of 50 ms. Five symmetric ($\Delta f/f_0$: 0.0; 0.1; 0.2; 0.3; 0.4) and two asymmetric notch conditions ($\Delta f/f_0$: 0.2|0.4; 0.4|0.2) were utilized, where Δf denotes the spacing between the inner noise edges and the signal frequency f_0 . The outside edges of the noise maskers were fixed at $\pm 0.8 \cdot f_0$.

Masked detection

The masking thresholds for 750-Hz tones at a fixed level of 65 dB SPL were measured with diotic as well as dichotic (uncorrelated) bandlimited white noise (50-1500 Hz). The tones of 500 ms duration were temporally centered in the 700 ms noise maskers. Maskers and tones were \cos^2 gated with ramp durations of 100 ms and 200 ms, respectively. The binaural masking level difference (MLD), $N_0S_0 - NuS_0$, is given as the difference between masking threshold for the diotic and dichotic condition.

Lateralization

Lateralization thresholds were measured for the same sinusoidal stimuli (750 Hz, 500 ms duration, 200 ms ramps) as in the masked detection task, at a level of 70 dB SPL. The tones were lateralized by introducing a pure phase delay (IPD) to one of the ears. The long 200 ms onset/offset ramps were chosen to minimize gating cues to lateralization and to make sure that the auditory system's *fine structure* processing was assessed. In addition to performing the lateralization task in quiet, three conditions with background noise interferers were measured. These interferers were continuous bandlimited white noises (50-1500 Hz) presented diotically 40 dB (diot40) or 10 dB (diot10) below the individual masking level, and presented dichotically (uncorrelated) 20 dB (dich20) below the individual masking level. The masking levels were estimated from the masked detection measurement.

All stimuli were generated in MATLAB and converted to analog signals using a 24-bit D/A converter (RME DIGI96/8). The sampling rate was 44.1 kHz for the speech recep-

tion measurement, 48 kHz for the masking experiments and 96 kHz for the lateralization task. The stimuli were presented in a double-walled sound-attenuating booth via Sennheiser HD580 headphones.

Procedures

Speech reception

The SRT was defined as the signal-to-noise ratio (SNR) leading to 50% correct identification of the individual words in the DANTALE II sentences. The noise level was kept constant at 65 dB SPL while the sentence level was varied adaptively. For each subject and noise condition, a mean threshold was calculated from at least two (but normally three) repetitions. Before data collection, the listeners were trained on four lists of 20 sentences each.

Frequency selectivity

A three-interval, three-alternative, forced-choice (3I-3AFC) weighted up-down method (Kaernbach, 1991) was applied to track the 75%-correct point on the psychometric function. Based on the assumption that auditory filters are output driven (Rosen *et al.*, 1998), the masker level was varied adaptively while the signal level was kept constant. A run was terminated after 14 reversals. Threshold was defined as the arithmetic mean of all masker levels following the 4th reversal. A nonlinear minimization routine was used to find the best-fitting rounded-exponential filter in the least-squares sense, assuming that the signal was detected by the filter with the best signal-to-noise ratio.

Masked detection

The same 3I-3AFC method as for the frequency selectivity measurement (including threshold estimation) was applied. Also here, the masker level was varied adaptively while the signal level was kept constant.

Lateralization

A two-interval, two-alternative, forced-choice (2I-2AFC) weighted up-down method was applied to track 75% correct lateralization. The first interval always contained the zero IPD reference tone (in the median plane) while the second interval contained the tone which was randomly lateralized to the left or right side. The background interferer was presented continuously during the run. The listeners were instructed to indicate the direction of motion. A run was terminated after 14 reversals and threshold was defined as the geometric mean of all IPDs following the 4th reversal. For each subject and interferer condition, a mean IPD threshold was calculated from at least three repetitions. All listeners performed more than 1000 lateralization judgements before actual data collection commenced.

RESULTS

Speech reception

All HI listeners showed larger SRTs than the NH listeners for all noise conditions (average increase: 3.5 dB for speech-shaped noise, 8.8 dB for modulated noise and 4.1 dB for lateralized noise). Also, the masking release due to modulation of the noise was significantly smaller (on average by 5.3 dB) for the HI listeners [one-way ANOVA: $F(1,12) = 45.2$; $p < 0.0001$]. However, they did not show a decreased *spatial* release from masking compared to the NH listeners. Despite self-reported difficulties with speech perception, no degraded performance was found for subject *kwf*(OD). Although *lvf*'s (OD) SRTs were slightly increased compared to the NH listeners, the decline observed was much smaller than that for the HI listeners.

Frequency selectivity

Figure 2(a) shows the average 10-dB bandwidths obtained for the NH and HI listeners, as well as the individual bandwidth estimates for the two OD subjects. The HI listeners showed significantly elevated bandwidths compared to the NH listeners [$F(1,26) = 10.6$; $p < 0.005$], by a factor of 1.2. However, the results varied considerably across the HI listeners. For subject *lvf*(OD), the left ear bandwidth was significantly increased (compared to NH), while for *kwf*, both ears exhibited elevated bandwidths. The asymmetry between left and right ear bandwidths is depicted in Fig. 2(b). While no significant difference of bandwidth asymmetry was found between HI and NH, both OD subjects showed larger asymmetries than the NH and all but one of the HI subjects.

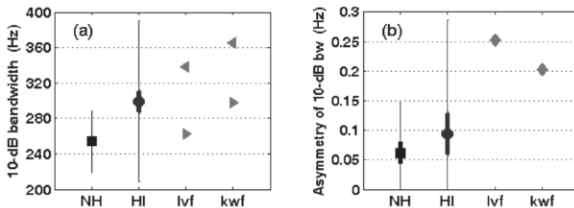


Fig. 2: (a) 10-dB bandwidth results; group mean for NH and HI listeners, single ear bandwidths for subjects *lvf* and *kwf* (<: left, >: right). (b) Asymmetry between left and right ear bandwidths, normalized to the individual's mean bandwidth. Thick error bars: standard error of group mean (for NH ↔ HI comparison), thin error bars: 2 standard deviations (for comparison of *lvf* and *kwf* with NH and HI).

Masked detection

In the masked detection task, the HI listeners showed significantly higher thresholds than the NH listeners, both for dichotic noise [$F(1,12) = 18.5$; $p = 0.001$] (Fig. 3a) and diotic noise [$F(1,12) = 4.5$; $p = 0.05$] (not shown here). Also the MLD (N_0S_0 - NuS_0) was significantly reduced [$F(1,12) = 6.6$; $p = 0.025$], as shown in Fig. 3(b). Both OD subjects performed significantly worse than NH on dichotic detection. However, their MLDs were not reduced significantly.

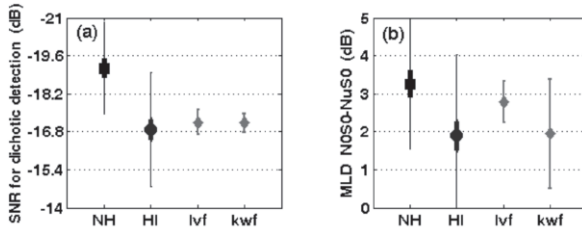


Fig. 3: (a) Masking threshold (SNR) for dichotic noise. (b) Masking release NOS0-NuS0. Errorbars for *lvf* and *kwf*: standard error of threshold estimate; others as in Fig. 2.

Lateralization

The results of the lateralization experiment are shown in Fig. 4. All statistics were performed on the \log_{10} of the IPD threshold values. As shown in the left panel, on average, lateralization performance for the NH listeners was significantly affected only in the presence of dichotic noise (dich20). Consistent with Lacher-Fougère and Demany (2005), a slight improvement was found for the low-level diotic noise (diot40) compared to the condition in quiet. The HI subjects *akm* and *lkf* showed more pronounced problems with lateralization than the other HI listeners. Therefore their results are excluded from the group mean and are discussed separately. The remaining HI listeners (denoted as HI*) showed significantly elevated lateralization thresholds (compared to NH) *only* in quiet [$F(1,10) = 6.3$; $p = 0.03$], as shown in Fig. 4(left). Particularly in the case of the low-level diotic noise (diot40), a significant improvement in performance can be seen for these listeners when compared to their performance in quiet [$F(1,10) = 6.3$; $p = 0.03$]. Fig. 4(right) shows individual results for the four listeners who performed markedly worse (group means from the left panel are plotted for comparison). Subject *akm* showed increased IPD thresholds, independent of interferer condition; subject *lkf* was not able to lateralize at all. Even at the maximum IPD of 90° her performance was not significantly different from chance level. Therefore, threshold IPD is plotted at 90° for this subject. The two OD subjects *lvf* and *kwf* showed markedly increased IPD thresholds as well. Both performed best in quiet but showed pronounced problems with lateralization (particularly subject *kwf*) in the presence of a noise interferer (even at a low level).

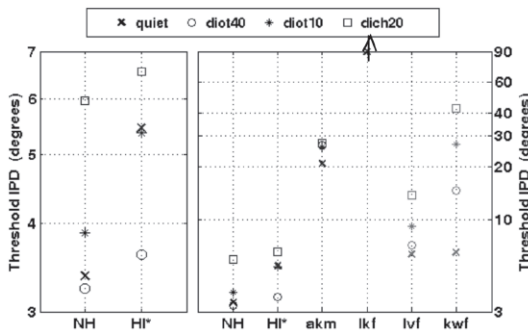


Fig. 4: Results of the lateralization threshold measurements.

Comparison of results across tests

Pearson sample correlations and 2-tailed p values were computed to examine the possible interrelations between frequency selectivity, masked detection, lateralization performance and speech reception within the group of HI listeners. No significant correlations between the frequency selectivity results and the other psychoacoustical measures were found in this study. Various significant correlations were found between lateralization performance and the masking thresholds for tones in dichotic noise as well as the MLD. Most prominent was the correlation between the IPD threshold in dichotic noise (dich20) and the corresponding tone masking threshold (SNR at detection) [$r = 0.86$; $p = 0.014$]. Regarding speech reception, the spatial release from masking (dichotic, lateralized noise condition) was found to correlate significantly with the masking threshold for tones in dichotic noise (SNR at detection) [$r = -0.92$; $p < 0.005$]. Otherwise, only marginal and thus inconclusive correlations between speech reception and lateralization performance were found. The correlations given above remained significant when controlling for hearing loss in terms of the PTA (average pure-tone threshold at 0.5, 1, 2 and 4 kHz) by means of partial correlation. This may be regarded as a consequence of the homogeneity of the HI group: the range between minimum and maximum PTA was restricted to 17.5 dB.

DISCUSSION

In the following, the results for the tests on frequency selectivity, binaural masked detection, binaural lateralization and speech reception are summarized and discussed. Frequency selectivity was found to be degraded in the HI and OD listeners, in spite of essentially normal audiometric thresholds up to 1 kHz (thus including the upper filter skirt). Deviant filter shapes seemed to reflect a possible underlying impairment, particularly for the two OD subjects. Also, the masking thresholds (SNR) at 750 Hz were significantly increased for the HI and OD subjects. Generally, stronger correlations with the other measures in this study were found for the threshold in dichotic noise (N_uS_0) than for the corresponding MLD (N_0S_0 - NuS_0). This is consistent with previous reports, that the MLD is generally less affected by hearing impairment than the dichotic threshold itself (Gabriel *et al.*, 1992). Our findings obtained in the lateralization experiment were not anticipated: for the majority of the HI listeners, significantly elevated IPD thresholds were found only in quiet but not in noise. Thus, for these listeners there was no indication of an increased vulnerability of binaural fine-structure processing to interfering noise. The finding of increased IPD thresholds in quiet despite normal audibility is in agreement with previous studies (e.g., Koehnke *et al.*, 1995). The results in noise could be explained as follows: for the tone level of 70 dB SPL, excitation is spread over a certain range of the basilar membrane, particularly for the HI listeners that exhibit reduced frequency selectivity. For the NH listeners, portions of the membrane corresponding to higher frequencies are likely to contribute to the lateralization judgement. However, the HI listeners might not be able to benefit from this information at higher frequencies since they fall into the region of hearing loss. On the contrary, if actually included into the decision process, information from defective units might have a detrimental effect on lateralization acuity. The additional noise would then confine the

excitation to the relevant region around 750 Hz and thus mask the deleterious spread at higher frequencies. Along with the fact that the diotic noise provides an additional, ongoing reference cue, this might explain the large improvement in lateralization seen for the HI listeners when going from quiet to low-level diotic noise. This hypothesis could be tested by repeating the experiment at a lower tone level where the effect of excitation spread should be expected to be much smaller. The correlations seen for the HI subjects between lateralization performance and the masking thresholds for tones in dichotic noise (as well as the MLD) could be indicative of a causal relation between binaural detection and lateralization. This relation seems to be robust regarding that the level of the noise interferer in the lateralization task was adjusted according to the individual's masking threshold. This choice of interferer level may also be partly responsible for the fact that no significant correlation was found between performance in lateralization and frequency selectivity. As a consequence, these two tests may be regarded as 'independent' measures of impairment factors (although caution has to be taken as absence of correlation does not imply independence). This can be illustrated with individual results among the HI listeners: while subject *lkf* showed the worst performance of all listeners in lateralization and frequency selectivity, subject *gaf* performed best in the former and second worst in the latter. Subject *akm* however, performed second worst in lateralization but best in frequency selectivity. In terms of speech reception, the HI listeners showed significantly degraded performance compared to the NH listeners. However, the essentially normal speech results for the OD subjects were not consistent with their problems in the other psychoacoustical tasks in this study. The redundancy of the full-spectrum sentences in connection with the SRT measure (instead of an intelligibility measure) might have been too high to track the influence which the identified psychophysical impairment factors (around 750 Hz) exerted on speech perception. This argument gets additional support from the finding that it was particularly the spatial masking release for speech which was correlated to the dichotic masking threshold at 750 Hz. As previously proposed by Levitt and Rabiner (1967), the binaural masking release for speech is primarily based on the interaural phase opposition at frequencies below about 500 Hz, suggesting that a correlation with our low-frequency factors is most likely found for this speech condition.

CONCLUSIONS

Performance on tests of frequency selectivity, binaural masked detection, binaural lateralization and speech reception was measured for listeners with normal hearing, impaired hearing and an obscure dysfunction. In spite of normal audiometric thresholds, significantly degraded performance in frequency selectivity, masked detection and lateralization was found for the listeners with impaired hearing and for the listeners with an obscure dysfunction. The SRT for full-spectrum speech seems to be insufficient to detect consequences of the basic impairment factors identified at low frequencies, such as frequency selectivity or fine-structure processing. For the hearing-impaired listeners no indication was found of an increased vulnerability of binaural fine-structure processing to interfering noise.

ACKNOWLEDGEMENTS

We wish to thank Graham Naylor, Steven Greenberg, Eric Thompson and our colleagues at the Centre for Applied Hearing Research for valuable and stimulating discussions.

REFERENCES

- Buss, E., Hall, J. W., and Grose, J. H. (2004). "Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss," *Ear. Hear.*, **25**(3), 242–250.
- Dreschler, W. A., and Plomp, R. (1985). "Relations between psychophysical data and speech perception for hearing-impaired subjects. II," *J. Acoust. Soc. Am.*, **78**(4), 1261–1270.
- Gabriel, K. J., Koehnke, J., and Colburn, H. S. (1992). "Frequency dependence of binaural performance in listeners with impaired binaural hearing," *J. Acoust. Soc. Am.*, **91**(1), 336–347.
- Horst, J. W. (1987). "Frequency discrimination of complex signals, frequency selectivity, and speech perception in hearing-impaired subjects," *J. Acoust. Soc. Am.*, **82**(3), 874–885.
- Kaernbach, C. (1991). "Simple adaptive testing with the weighted up-down method," *Percept. Psychophys.*, **49**(3), 227–229.
- Koehnke, J., Culotta, C. P., Hawley, M. L., and Colburn, H. S. (1995). "Effects of reference interaural time and intensity differences on binaural performance in listeners with normal and impaired hearing," *Ear. Hear.*, **16**(4), 331–353.
- Lacher-Fougère, S., and Demany, L. (2005). "Consequences of cochlear damage for the detection of interaural phase differences," *J. Acoust. Soc. Am.*, **118**(4), 2519–2526.
- Levitt, H., and Rabiner, L. R. (1967). "Binaural release from masking for speech and gain in intelligibility," *J. Acoust. Soc. Am.*, **42**(3), 601–608.
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. J. (2006). "Speech perception problems of the hearing impaired reflect inability to use temporal fine structure," *Proc. Natl. Acad. Sci. USA*, **103**(49):18866–18869.
- Moore, B. C. (1996). "Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids," *Ear. Hear.*, **17**(2), 133–161.
- Patterson, R. D., and Nimmo-Smith, I. (1980). "Off-frequency listening and auditory-filter asymmetry," *J. Acoust. Soc. Am.*, **67**(1), 229–245.
- Rosen, S., Baker, R. J., and Darling, A. (1998). "Auditory filter nonlinearity at 2 kHz in normal hearing listeners," *J. Acoust. Soc. Am.*, **103**(5 Pt 1), 2539–2550.
- Santurette, S., and Dau, T. (2007). "Binaural pitch perception in normal-hearing and hearing-impaired listeners," *Hear. Res.*, **223**(1-2), 29–47.
- Wagener, K., Jøsvassen, J. L., and Ardenkjaer, R. (2003). "Design, optimization and evaluation of a Danish sentence test in noise," *Int. J. Audiol.*, **42**(1), 10–17.

