Spatio-temporal integration of speech reflections in listeners with sensorineural hearing loss

Jan Rennies[1], Anna Warzybok[2], Birger Kollmeier[1,2], Thomas Brand[2]

[1] *Fraunhofer IDMT, Hearing, Speech and Audio Technology & Cluster of Excellence Hearing4all, Oldenburg, Germany, E-Mail: jan.rennies@idmt.fraunhofer.de*
[2] *University of Oldenburg, Medical Physics and Acoustics &Cluster of Excellence Hearing4all, Oldenburg, Germany*

Introduction

In rooms speech is reflected at boundaries and objects and superimposes with the direct sound. This creates a complex pattern of temporally delayed, spectrally modified and spatially distributed copies of the direct sound. These reflection patterns are reflected in the room impulse response (RIR). For speech intelligibility in rooms, early reflections are typically assumed to be beneficial while late reflections are considered to be detrimental. This concept was recently challenged in a recent study employing binaural RIRs with systematically varied interaural phase differences (IPDs) of the direct sound and a variable number of reflections delayed by up to 200 ms [1]. Speech recognition thresholds (SRTs) in stationary noise measured in normal-hearing listeners showed that, when IPD favored late RIR components, listeners appeared to be capable of focusing on these components rather than on the precedent direct sound, which is at odds with common room acoustical measures. The goal of this study was to extend these results to listeners with sensorineural hearing loss to explore to what degree they can exploit IPD cues under the same conditions.

Methods

Subjects

Fourteen subjects with symmetrical, sloping hearing loss between 45 and 84 y (Ø 74.5 y) participated in this study (see Figure 1). Their PTA_4 (average pure-tone audiogram across 0.5, 1, 2, and 4 kHz) ranged from 26 to 58 dB HL (Ø 43 dB HL). The average hearing loss across 0.125, 0.25, 0.5, 0.7 and 1 kHz ($PTA_{\leq 1k}$) was between 10 and 44 dB HL (Ø 27 dB HL), i.e., some subjects had normal hearing in the frequency range primarily used for IPD processing. All subjects received individual linear amplification according to NAL-RP to compensate for audibility loss.

Procedure

SRTs were measured using the Oldenburg sentence test [2]. The target speech was convolved with artificially created RIRs consisting of direct sound and a variable number of reflections, which had the same amplitude as the direct sound. The IPD of the direct sound (D) or the reflections (R) was systematically manipulated by inverting the corresponding RIR components at the left ear. Diotic, stationary speech-shaped noise at a fixed level of 65 dB SPL was used as masker (N_0). SRTs were measured using an adaptive procedure, which converges to the point of 50% speech intelligibility [2].

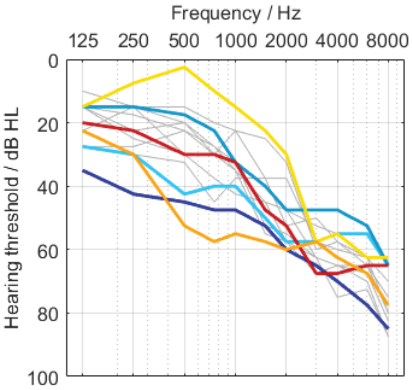

Figure 1: Average audiograms across ears of the 14 participants. Blueish colors represent the 3 subjects with the worst performance in the reference condition, reddish/yellow colors represent the 3 best subjects. Intermediate subjects are shown in gray.

Results

Figure 1 shows SRTs measured in conditions with direct sound only, i.e., when no reflections had to be integrated with the direct sound. The left-most data points in each panel show the baseline condition (D_0N_0), in which no binaural information was present either (no temporal integration, no binaural processing). Here and in the following, the data of hearing-impaired listeners (left panel) are compared to the data of the normal-hearing listeners (right panel). The data in the baseline condition indicated that, in general, aided hearing-impaired subjects could do the simple speech recognition task very well (close to normal performance). There was some interindividual variability (minimum and maximum SRTs differed by about 4 dB), which could not be explained by audiogram-derived measures (PTA_4: $R^2<0.01$; $PTA_{\leq 1k}$: $R^2=0.01$). In this condition, the best [worst] three subjects are indicated by reddish/yellow [blueish] symbols, and the same colors are used in all subsequent figures to track the performance of these subjects across conditions (see also corresponding audiograms in Figure 1). The second and third group of data points in Figure 2 show SRTs for simple binaural processing, in which IPD cues were available (either as D_π or N_π), but no temporal integration of reflections was required. These conditions showed that all hearing-impaired subjects benefited from IPD advantages to some degree, although the average benefit was slightly smaller than for normal-hearing listeners. Interindividual variance increased in these binaural conditions.

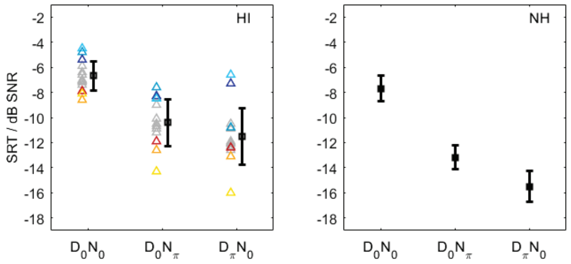

Figure 2: SRTs of hearing-impaired (left) and normal-hearing subjects (right) in conditions without temporal integration.

There was a large interindividual variance in IPD advantage across subjects (see Figure 3). The IPD-benefit for $D_\pi N_0$ could be partly explained ($p<0.05$) by the (low-frequency) hearing loss ($PTA_{\leq 1k}$: $R^2=0.57$; PTA_4: $R^2=0.43$).

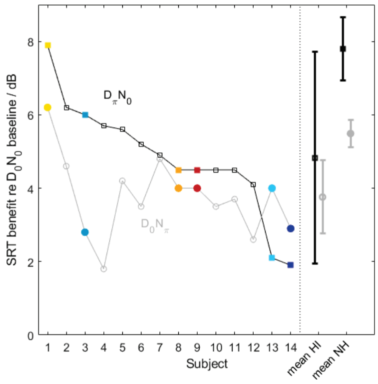

Figure 3: IPD-advantage in the conditions without temporal integration of speech reflections.

Figure 4 shows SRTs measured in diotic conditions, in which temporal processing was required because the RIRs consisted of direct sound and a single, equally strong reflection with varying delay. It was found that SRTs increased with delay similar to what had been observed for normal-hearing listeners. Some subjects could cope with increasing delay as well as normal-hearing subjects, while other subjects showed a much larger SRT increase at the longest delay of 200 ms. On average, the SRT increase was about 3 dB larger. The individual delay deficit at 200 ms could not be explained by hearing loss, but the SRTs at a delay of 200 ms could be explained in part by baseline SRTs ($R^2=0.56$, $p=0.002$).

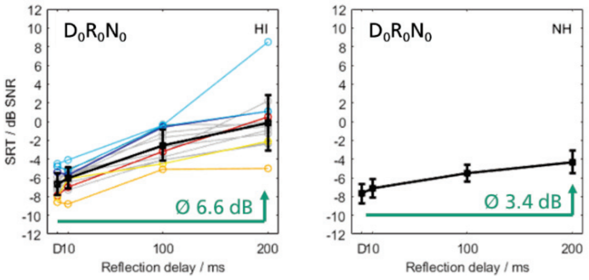

Figure 4: SRTs of hearing-impaired (left) and normal-hearing subjects (right) in conditions with a single reflection.

Figure 5 shows SRTs of conditions in which both temporal and processing was required. Here, the same reflection delays as before were used, but the reflection had a π-IPD and thus carried binaural information that could be exploited in diotic noise. In [1] it was found that SRTs decreased considerably when adding a reflection with π-IPD to the direct sound, and that the SRT decrease was independent of reflection delay. In other words, normal-hearing listeners appeared to be able to focus on the late reflection even when it was strongly delayed. The data of the hearing-impaired data of the present study showed that all subjects could benefit from an IPD when the reflection delay was short (10 ms). However, at a delay of 200 ms, there was a large variance, which could partly be predicted by the individual IPD benefit ($R^2=0.55$, see Figure 3) or by the diotic reference SRTs ($R^2=0.61$).

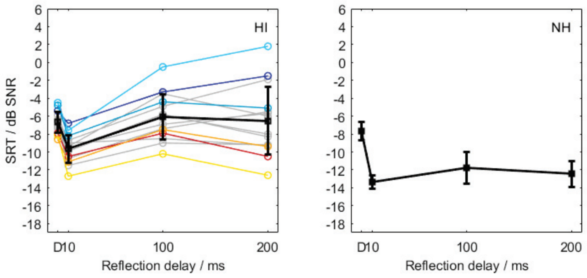

Figure 5: SRTs of hearing-impaired (left) and normal-hearing subjects (right) in conditions with a single reflection which carried an IPD advantage in diotic noise.

Conclusions

The present group of (aided) hearing-impaired listeners had normal or close-to-normal SRTs in the baseline condition and normal or smaller-than-normal residual binaural hearing benefit. They could also generally cope with a single (even very late) reflection, although the interindividual variability was large for large delays. However, their (residual) capability to use IPD-information did not automatically mean that subjects could extract information from a late reflection in the way normal-hearing subjects can: Some hearing-impaired subjects (the “generally better” ones) were close-to-normal even in complex spatio-temporal conditions, but this was not true for all subjects with good low-frequency hearing. Simple audiogram measures were not good predictors of SRTs in complex spatio-temporal conditions, indicating a need for more advanced predictions models to better understand the complex interaction between hearing loss and spatio-temporal integration revealed by the data of this study.

Acknowledgements

This study was funded by the Klaus Tschira Stiftung gGmbH and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 352015383 – SFB 1330 A1.

References

[1] Rennies, J., Warzybok, An. Brand, T. & Kollmeier, B. (2019). Measurement and prediction of binaural-temporal integration of speech reflections, Trends Hear. 23, 1-22.

[2] Wagener, K., Kühnel, V. & Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests (Development and evaluation of a German sentence test I: Design of the Oldenburg sentence test). Z. Audiol. 38(1), 4-15.