Binaural listening effort in noise and reverberation

Jan Rennies[1,2], Gerald Kidd, Jr.[1]

[1]Speech, Language and Hearing Sciences and Hearing Research Center, Boston University, Boston, MA, Email: jrennies@bu.edu; gkidd@bu.edu

[2]Fraunhofer IDMT, Hearing, Speech and Audio Technology, Oldenburg, Germany,

Introduction

Speech intelligibility is classically measured as percentage of correctly understood syllables, words, or sentences, or as speech reception thresholds (SRTs), i.e., the SNR at which half of the speech items can be understood correctly. Various studies have addressed factors affecting speech intelligibility and the role of binaural hearing, and the importance of interaural level and time differences (ILDs/ITDs) have been investigated in depth. However, speech intelligibility often approaches ceiling performance already in very unfavorable conditions, e.g., at negative SNRs, which are not necessarily representative for everyday listening conditions. The ecological validity of measuring speech intelligibility is hence limited to some degree.

In contrast, listening effort, defined here as the perceived effort associated with recognizing target speech, can be measured across a wide range of SNRs and other detrimental factors such as reverberation [1].

Despite this advantage and a recent trend to employ listening effort for scientific evaluations in audiology (e.g., [2, 3]) and speech enhancement (e.g., [4]), relatively few studies have addressed listening effort in binaural listening conditions. The research question of this study was therefore whether there is spatial release from listening effort in conditions with systematically varied binaural masking, especially in conditions in which speech intelligibility is at ceiling.

Methods

Seven subjects (4 female / 3 male) with normal hearing (age 18-30 years) participated in this pilot study. English target speech material from the matrix sentence test [5] was presented at a fixed level of 55 dB SPL. A single stationary speech-shaped noise source with the same long-term spectrum as the target speech was used as masker (one condition with two noise sources was also used). The level of the noise was varied to produce desired SNRs. Speech and masker were convolved with binaural head-related impulse responses recorded in an anechoic room to generate the desired spatial conditions. The target speech was always presented from the front (S_0), while the masker was presented from different azimuthal directions (N_0, N_{45}, N_{90}, N_{135}, N_{180}, and $N_{-45}N_{135}$). Stimuli were presented to the subjects over Sennheiser HD280 pro headphones

Listening effort was measured using categorical listening effort scaling, i.e., subjects heard the stimuli and rated the listening effort on a scale (Figure 1) with seven named categories ranging from "no effort" (corresponding to 1 effort scaling categorical unit, ESCU) to "extreme effort" (13 ESCU), and six unnamed categories in-between the named ones. The additional category "only noise" was provided for the case that subjects could not hear the target talker at all. Speech and masker were mixed at fixed SNRs (as measured at the right ear) from -25 to +10 dB, and each combination of noise azimuth and SNR was assessed four times by each subject. Psychometric functions were obtained for each subject and condition by fitting model functions consisting of one or two straight lines to the categorical ratings as a function of SNR [3].

Figure 1: Categorical listening effort scale.

Results

Symbols in Figure 2 represent the median listening effort ratings across subjects, errorbars indicate interquartile ranges. Lines show the model functions fitted to the data. For all masking conditions, listening effort decreased with increasing SNR as expected. There was no difference in listening effort in the N_0, N_{180}, $N_{-45}N_{135}$ conditions (dashed lines), i.e., the conditions with (rather) symmetric maskers at the two ears. In these conditions, listening effort varied from "extreme effort" to "no effort" in the SNR range from about -15 to +14 dB.

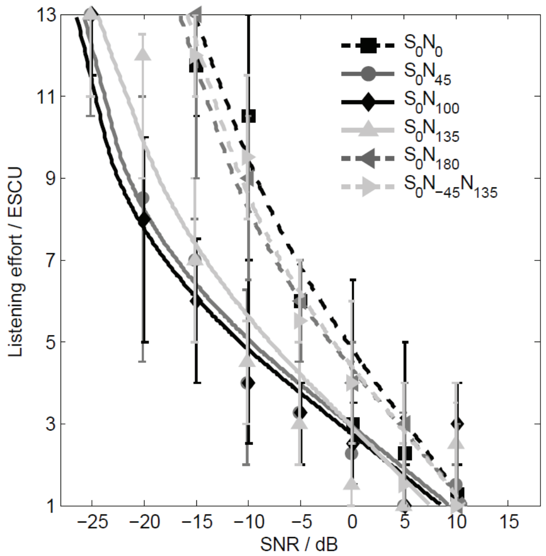

Figure 2: Listening effort as a function of SNR. Symbols show inter-individual medians and interquartile ranges, lines represent model functions fitted to the median data across subjects.

In conditions with a lateral masker (N_{45}, N_{90}, N_{135}), listening effort was considerably lower at low SNRs. At the highest effort rating of 13 ESCU, the difference in SNR compared to symmetric masking conditions was about 10 dB. This spatial advantage decreased with increasing SNR, and was absent at very favorable SNRs.

The spatial release from listening effort is further illustrated in Figure 3, which shows iso-listening-effort contours derived from the model functions. Each curve shows the SNRs resulting in a constant listening effort rating as a function of noise azimuth. The condition with two maskers is shown on the right-hand side. Separate Friedman tests for each ESCU value with independent variable SNR and factor noise azimuth show a significant main effect of noise azimuth for all ESCU values except 1 and 2 at a corrected significance level of 0.05/13.

Discussion

Data for co-located and symmetric maskers are in line with previous studies showing that listening effort decreases with increasing SNR. The present data show that spatially separating a single noise source from a frontal target is beneficial in terms of listening effort over a large range of SNRs. At very high SNRs listening effort is very low in all conditions and no benefit due to spatial unmasking is observed. At intermediate SNRs, spatial release from listening effort is significant also in well inside the range of positive SNRs, in which speech intelligibility is at ceiling [1]. Further data on binaural listening effort in a reverberant room (not shown in this paper) suggest that – in analogy to the well-known decrease in spatial unmasking for speech intelligibility – reverberation also reduces the binaural advantage in listening effort. Listening effort scaling therefore seems to be a valuable tool for evaluating real-life, binaural listening conditions.

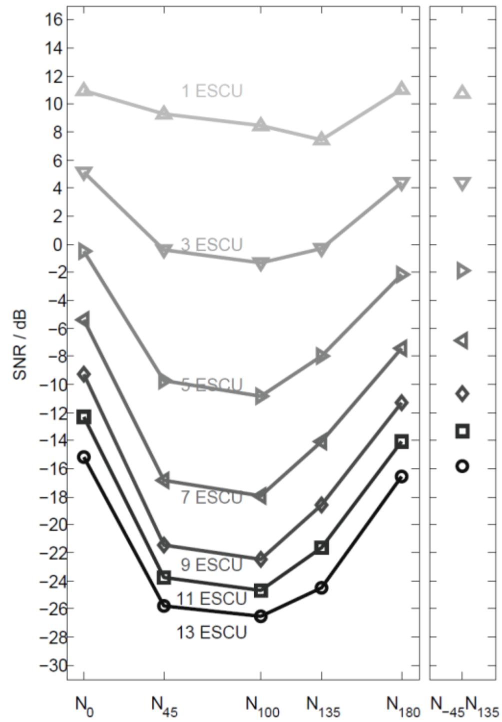

Figure 3: Iso-listening-effort contours as a function of noise azimuth.

Acknowledgements

This study was supported by the German Research Foundation (DFG) Grant No. RE 4160/1-1, and the National Institutes of Health–National Institute on Deafness and other Communication Disorders Grant No. R01 DC04545.

References

[1] Rennies, J., Schepker, H., Holube, I., and Kollmeier, B. (2014). “Listening effort and speech intelligibility in listening situations affected by noise and reverberation,” J. Acoust. Soc. Am. 136, 2642-2653.

[2] van Schoonhoven, J., Schulte, M., Boymans, M., Wagener, K. C., Dreschler, W. A., and Kollmeier, B. (2016). “Selecting appropriate tests to assess the benefits of bilateral amplification with hearing aids,” Trends Hear. 20, 1–16.

[3] Krueger, M., Schulte, M., Brand, T., and Holube, I. (2017a). “Development of an adaptive scaling method for subjective listening effort,” J. Acoust. Soc. Am. 141, 480-4693.

[4] Pusch, A., Rennies, J., Schepker, H. and Doclo, S. (2018). “Höranstrengung als Messverfahren zur Evaluation von Near-end listening enhancement Algorithmen,” Fortschritte der Akustik - DAGA 2018, Munich, Germany.

[5] Wagener, K., Brand, T., and Kollmeier, B. (1999). “Entwicklung und Evaluation eines Satztests für die deutsche Sprache III: Evaluation des Oldenburger Satztests (Development and evaluation of a German sentence test III: Evaluation of the Oldenburg sentence test).” Z. Audiol. 38, 86-95.

[6] Rennies, J., Brand, T., and Kollmeier, B. (2011a). “Prediction of the influence of reverberation on binaural speech intelligibility in noise and in quiet,” J. Acoust. Soc. Am. 130, 2999-3012.