PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Speech intelligibility in a realistic virtual sound environment

Naim Mansour[1], Marton Marschall[1], Adam Westermann[2], Tobias May[1], Torsten Dau[1]

[1] Hearing Systems Group, Technical University of Denmark (DTU)

[2] Widex A/S, Lynge, Denmark

ABSTRACT

Employing realistic, yet controlled sound scenarios for the evaluation of hearing aid (HA) algorithms in a virtual sound environment (VSE) has the potential to improve HA users' real-world listening experience and performance. In the present study, so-called critical sound scenarios (CSS) were defined as acoustic scenes that HA users experience as important, difficult and common, and selected through ecological momentary assessment (EMA), inspired by results from recent studies. The scenarios were acquired in real scenes using a spherical microphone array, and reproduced in an acoustically valid way inside the VSE. A speech intelligibility task using the Danish Hearing In Noise Test (HINT) was implemented in one of the scenes to measure speech reception thresholds (SRT) for normal-hearing and hearing-impaired subjects. SRTs obtained for the realistic VSE background noise were found to be higher than corresponding ones reported in headphone-based studies, potentially as a result of the increased cognitive effort required to separate target speech from the more realistic interfering speech background.

Keywords: Speech intelligibility, Experimental realism, Spatial audio

1. INTRODUCTION

The tremendous sophistication of auditory perception allows humans to make sense of and communicate in vastly complex acoustic scenes and situations. Consequently, impairments to this system can have a negative impact on social interactions of those affected (1). The development of increasingly complex hearing devices over the past few decades has been aimed at partially compensating for these impairments. Complex algorithms inside hearing aids (HA) control the gain of incoming sound in a time- and frequency-dependent way. In order to achieve maximum performance when worn by users in their daily lives, it makes sense to evaluate and validate HAs and HA algorithms in conditions that mimic those of the real world. Moreover, results of psychoacoustic listening tasks involving normal-hearing (NH) and (aided) hearing-impaired ((A)HI) subjects might be affected by the increased experimental realism these conditions can provide.

The design of such psychoacoustic experiments can be viewed in light of a trade-off between control and realism. Experimental outcome measures benefit from being clear and consistent, with their paradigms accurately described and replicable. In a listening task, this implies having control over the precise acoustic conditions and variabilities in the stimuli. At the same time, a design that does not incorporate the complexity of real-world conditions might not produce generalizable results.

Specifically, a listening task that uses headphones to capture speech-in-noise performance can spatialize stimuli using head-related-transfer functions, but this approach is coarse and its accuracy is limited by many factors, including headphone placement and the quality of said transfer functions. In addition, accounting for head movements and fitting a HA is cumbersome in practice. Quadrophonic loudspeaker setups can be employed to spatialize the noise maskers, alleviating these problems, but they generally still do not produce signals with realistic spatial properties. In contrast, conducting the task in the real world would be perfectly realistic, yet real-world conditions tend to be poorly controlled and highly variable. Therefore, an experimental setup that trades off control for realism in an optimal way has the potential to result in well-controlled, yet ecologically valid findings.

In addition to a controlled and realistic reproduction, the type of source signals (target as well as masker) can affect experimental realism as well. Typically used speech-shaped noise maskers tend to be quite dissimilar to an actual multi-talker acoustic scene. They do not contain low-frequency modulation masking components, have shallower time-varying dynamics and lack intelligible components which could lead to informational masking (2).

Simulating spatialized room-acoustic maskers (such as those obtained through ODEON software (3)) has been shown to increase speech reception thresholds (4), but might still be limited in level of realism by the number and complexity of sources that can be simulated. In order to maximize realism, a controlled, in-situ spatial recording of a realistic sound scenario is likely to be even more representative, retaining auditory cues up to the limitations of the recording setup.

Finally, any recorded sound scenario would benefit in terms of its ecological validity from having been selected based on actual HA user input. The use of ecological momentary assessment (EMA) data captured in the real world has become increasingly prevalent in attempts to qualitatively and quantitatively determine which scenes HA users experience (5). Thus, choosing scenarios that are aligned with these experiences is a logical way to further improve experimental realism (6).

To evaluate the specific speech-in-noise paradigm of speech intelligibility (SI), the Hearing In Noise Test (HINT) has been widely used, in various acoustic conditions. This test, along with other approaches, such as the matrix-based DANTALE II (7), results in NH SRTs that are well below zero, at around -2.5dB in the Danish HINT (8), -3.9dB in a multi-lingual HINT pooling (9), and down to -8dB for the DANTALE II.

In contrast, research trying to categorize real-world signal-to-noise ratios (SNRs) has consistently found a substantially higher range of SNR values in the majority of sound scenarios (10). Thus, when relating SI test scores to intelligibility in a realistic scenario, headphone-based tasks using artificial stimuli might lead to an underestimation of the difficulties both NH and HI people experience.

The speech-in-noise experiment considered in the present study attempts to address some of the mentioned limitations, by designing an ecologically valid SI task based on the Danish HINT. At the core of the study lie two hypotheses. The first one states that a speech intelligibility paradigm that employs realistic, spatialized stimuli will produce SRTs that are significantly higher compared to those obtained with an artificial approach. In the second hypothesis, those obtained SRTs are proposed to be consistent with SNRs reported in the real world.

A set of critical sound scenarios (CSS) was selected based on a novel categorization of existing HA user EMA data (Figure 1a). These scenarios were recorded in-situ with a spherical microphone array (Figure 1b). A 64-channel, fully spherical loudspeaker array inside an anechoic enclosure (Figure 1c), capable of rendering arbitrary three-dimensional sound fields at its center through Higher-Order Ambisonic reproduction techniques (HOA, 11), was used to reproduce these recordings as a so-called Virtual Sound Environment (VSE). Finally, all components were combined into the design of a speech intelligibility task for NH and HI subjects (Figure 1d-e) based on the Danish HINT.

a. Selection b. Acquisition c. Reproduction d. SI for NH e. SI for HI

Figure 1 – Design stages of the VSE SI task

2. METHOD

2.1 Sound scenario selection

To establish a high level of ecological validity in the speech intelligibility task, the acoustic background stimuli were selected based on EMA data from 281 field reports by HA users, collected by Smeds et al. (12). Their data were categorized into so-called critical sound scenarios (CSSs) in a novel way, based on the orthogonal division between three reported EMA metrics: importance, occurrence and difficulty (Fig. 2). Combinations of these three parameters seem to separate different real-life situations very accurately. Some examples in Figure 2 illustrate this.

Using CSSs in the "sweet spot" (the subset of scenarios that are important, difficult and common), as backgrounds for the SI task was intended and assumed to reflect conditions in which HA users struggle. Important, difficult and less common CSSs are arguably also of interest for selection, since they still reflect challenging conditions for the HA to perform in. Three sweet spot scenarios were selected: a public lunch, a small festive event and an office meeting. The speech intelligibility task in this study was implemented only for the office meeting scenario.

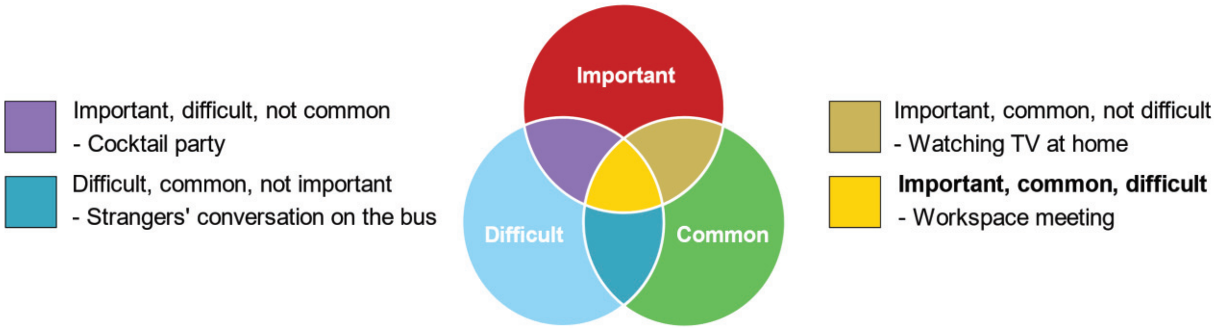
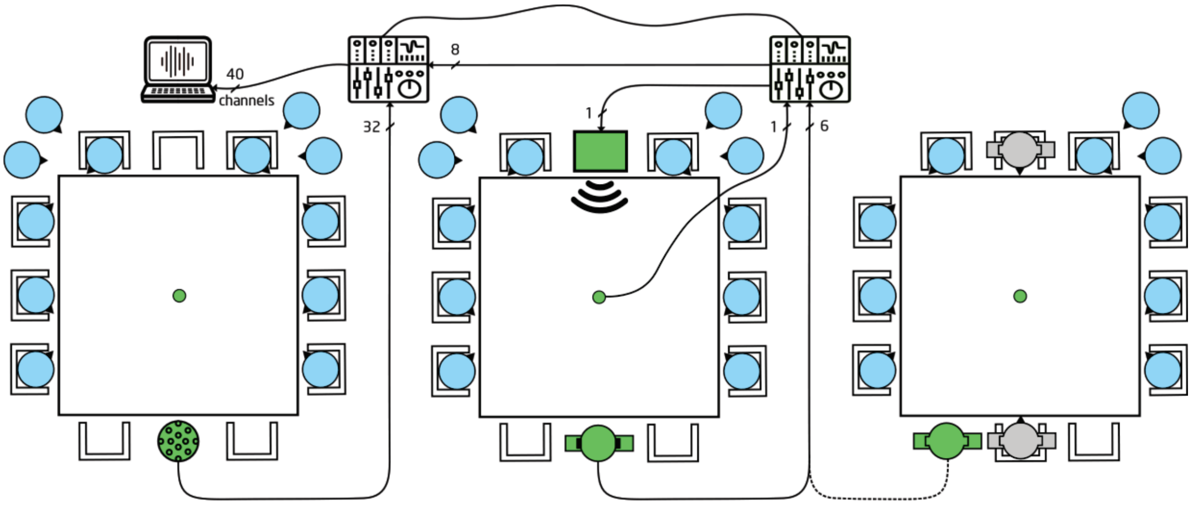


Figure 2 – Critical Sound Scenario framework

2.2 Sound scenario acquisition

The selected sound scenarios were captured in the real world with an em32 Eigenmike spherical microphone array capable of 4th HOA recording (13). In addition, a Knowles Electronic Manikin for Acoustic Research (KEMAR) equipped with HA shells was used to capture binaural ear and HA signals. An overhanging omnidirectional microphone measured overall background levels. Figure 3 illustrates the scenario acquisition for the office meeting scenario.



a. Background scene recording b. Impulse response recording c. Conversation in noise recording

Figure 3 – Office meeting scenario acquisition stages

Spatial background recordings were obtained with the EigenMike and KEMAR in the listener position, while 12 participants, seated and standing in the office meeting room, conversed as illustrated in Figure 3a. Impulse responses were captured from the target position using a mounted loudspeaker producing exponential sweeps, while all participants remained in place to avoid altering the room reverberation (Figure 3b). This was done to later allow virtual embedding of the target speech material into the room. In addition to these recordings, the scene was captured by the right ear of the KEMAR positioned to the left of a participant sitting in the listener position, conversing with a partner in the target position across the table (Figure 3c). This last stage allowed the capture of the background scene during a target-listener conversation. At each stage, the signals were clock-synchronized to single-sample precision at a 48kHz sample rate.

2.3 Sound scenario reproduction

The spherical microphone recordings obtained during acquisition were encoded from the 32 raw input channels to a 25-channel Ambisonic 4^{th}-order HOA format. These HOA signals were then rendered on a 64-channel spherical loudspeaker array (Figure 5b) using dual-band (basic, max-rE) decoding, with a crossover frequency of 2400 Hz. HOA auralization was chosen for its physically faithful rendering of sound fields in the sweet spot at the center of the array, ensuring their usability for HAs as well as human ears.

To guarantee that the overall root-mean-squared (RMS) level would not fluctuate too widely throughout the reproduction, an additional procedure was designed to select and combine portions of the auralized recording, based on level estimates derived from the overhanging reference microphone.

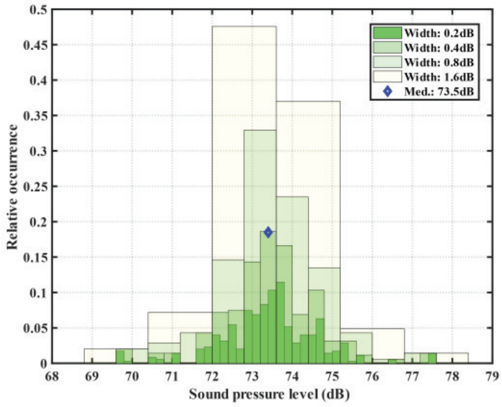


Figure 4 – Recorded background level distribution

The 10-minute-long reference microphone recording was segmented into frames of 5s (with 80% overlap). Level values in dB were derived for each frame, and the level difference calculated between consecutive frames. The upper and lower boundaries for allowed level differences were respectively set to the 5th and 95th percentile of the difference distribution. This allowed the segment indices of consecutive frames within these boundaries to be retained, ordered by decreasing length (from 25s to about 7s), and resulted in a 2.5 min. recording.

Next, the 64-channel synchronized spherical recording was decomposed into frames of the same length and overlap, and the portions corresponding to the selected reference segment indices were extracted. These frame segments were then individually calibrated to a fixed target level using a B&K 2669 free-field microphone, before being concatenated through windowed cross-fading with 1s overlap. The frame pairs at the cross-fading boundaries were calibrated in the same way, to ensure smooth transitions between segments. Finally, the resulting 2-min., 64-channel background recording was calibrated binaurally inside the loudspeaker array, using a B&K Type 4128 Head and Torso Simulator (HATS) with ear canals. As a result of this approach, the obtained background signal retained its intelligible properties despite being level-equalized in a time-sensitive way.

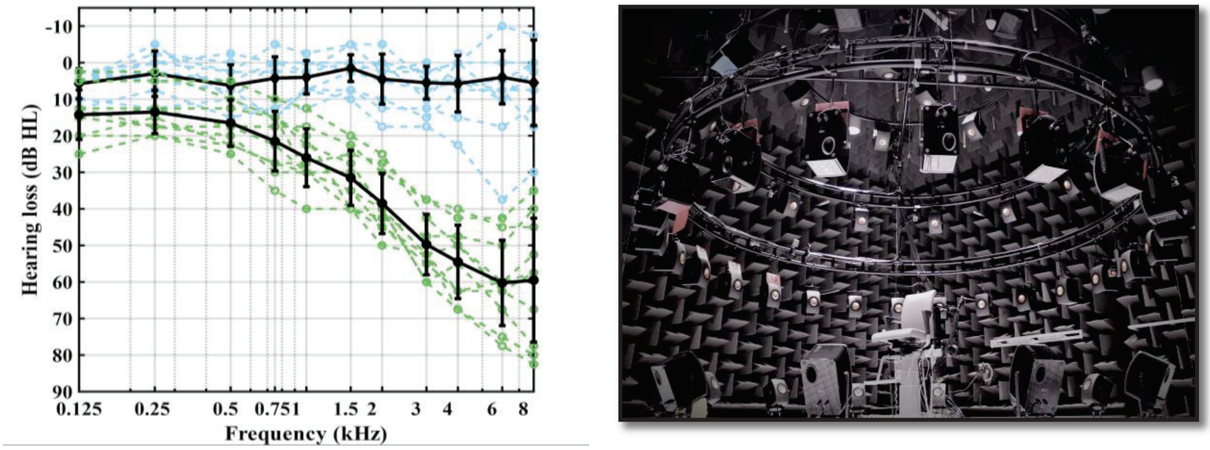
The target sound pressure level for the sound scenario, 73.5dB, was chosen as the median value of the noise level distribution measured during the third phase of acquisition. This distribution was calculated by selecting noise-only portions of the recording through the use of an energy-based voice-activity detector (14), and shown in Figure 4 as a nested histogram (with increasing bin widths of 0.2, 0.4, 0.8 and 1.6dB). The background noise level in the classical Danish HINT study (8) was set to values between 65 and 75dB SPL.

2.4 Speech intelligibility assessment

The testing paradigm to evaluate speech intelligibility was kept exactly the same as that of the classical Danish HINT. In its classical implementation, the well-established HINT uses brief, mundane target sentences superimposed on speech-shaped quasi-stationary noise, presented diotically over headphones in an adaptive 1-up-1-down sentence-based scoring procedure. The non-stimuli-related specifics of this procedure, as well as the SRT calculation method and program interface were left unaltered. To embed the 60 training and 200 test speech sentences in the spherical office meeting background, they were convolved with impulse responses measured between the target and listener positions by the spherical microphone array. Each sentence was individually calibrated to retain the same unique level as that of the single-channel version, referred to the sweet spot of the loudspeaker array in the AudioVisual Immersion Lab (AVIL) at DTU. Throughout the test, different SNRs were accomplished by varying speech levels with respect to the constant, continuously looped background.

10 NH and 10 HI subjects participated in the experiment. The NH subjects had a four-frequency average hearing loss of at most 15dB, while the HI ones had an average sloping mild (N2) to moderate (N3) hearing loss (15). Figure 5a shows the individual mean left-right as well as the overall mean audiograms with standard deviation for both groups. The NH/HI median age was 28/70 years old, with standard deviations of 16.6/5.3 years. All HI subjects had a discrimination performance of at least 92% for both ears, and a maximum left-right-ear difference of at most 10 dB for all frequencies.

The N2-N3 HI target criterion was intended to produce a homogenous and ecologically representative listener group with a common, but not overly severe hearing loss. Symmetrical hearing loss was required to avoid effects of better-ear listening in a task that allows for head movement.



a. Audiograms of NH and HI subjects

b. 64-loudspeaker array for HOA reproduction

Figure 5 – Audiograms & loudspeaker array in the AudioVisual Immersion Lab (AVIL)

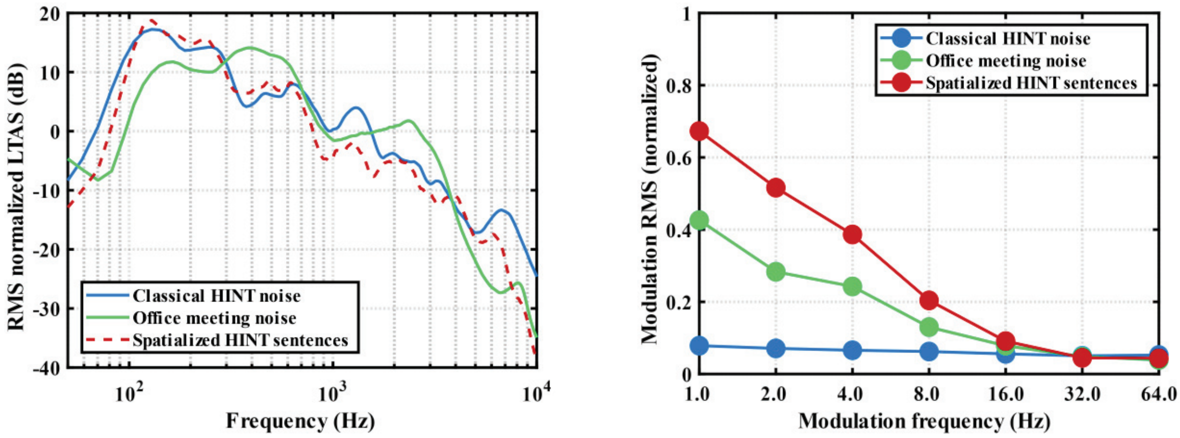
One training round and two evaluation rounds of 20 sentences each were carried out to determine SRTs per subject in the office meeting scenario background. A target-sentence-only training round was used to ascertain audibility, in accordance with the HINT procedure. The order of used speech material was randomized by arranging the individual test list orders in a Latin square, and an SRT score was determined for each subject as the average of the two evaluated SRTs. The reference HINT study included 16 NH and HI subjects, and SRTs values were determined per subject by averaging evaluated SRTs over 10 test lists. Figure 5b shows the AVIL loudspeaker array used in the experiment.

3. RESULTS

3.1 Acoustic properties of the background

Figure 6a and 6b respectively show the long-term average spectrum (LTAS) up to the Nyquist frequency, and the modulation power spectrum from 1Hz exponentially up to 64Hz, for the classical HINT background noise, the office meeting background noise, and a concatenation of all two hundred spatialized HINT test sentences.

The LTAS were normalized to have the same broadband RMS, and the modulation power spectra were obtained by normalizing calculated modulation power within a band by its respective bandwidth, as well as the power in the DC component (16). The auralized HINT sentences were recorded with the HATS inside the loudspeaker array (with the left ear spectra shown here). Several observations can be made. The LTAS and modulation spectrum of the classical HINT noise represent its quasi-stationary nature and lack of significant low-frequency modulation power. The modulation power is mostly constant for all considered modulation frequencies, resulting in a flat spectral representation.



a. Long-term average spectra

c. Modulation spectra

Figure 6 – Spectral properties of the SI stimuli

The situation is very different for the office meeting noise. The LTAS is largely similar in terms of its speech-like spectral shape, though containing fewer prominent peaks, and including the effect of the room. On the other hand, the modulation spectrum now has significant power at modulation frequencies up to 8Hz. The modulation spectrum of the office meeting background noise bears a greater resemblance to the auralized HINT sentences than to the classical HINT noise. The LTAS for the HINT sentences is naturally more similar to that of the classical noise, since this noise was originally constructed from the average power spectrum of one hundred HINT sentences.

Thus, the classical HINT noise takes the averaged spectral characteristics of the speech stimuli into account, but not the low-frequency modulations. The office background noise clearly includes these modulations, like the HINT sentences, while its LTAS resembles that of the classical HINT noise.

3.2 Speech reception thresholds

Fig. 7 shows individual SRTs as well as their median, mean and $25^{th}/75^{th}$ percentiles for the NH and HI subjects in both the classical, reference HINT study (Ref. - left) and the current study (VSE - right). The means/median values for NH_{ref} and HI_{ref} subjects are -2.52/-2.49dB and 0.09/0.25dB, respectively. Those for NH_{VSE} and HI_{VSE} are -1.17/-1.32dB and 1.54/1.58dB. The NH interquartile range is larger for the VSE condition compared to the reference condition, while the situation is reversed for the HI.

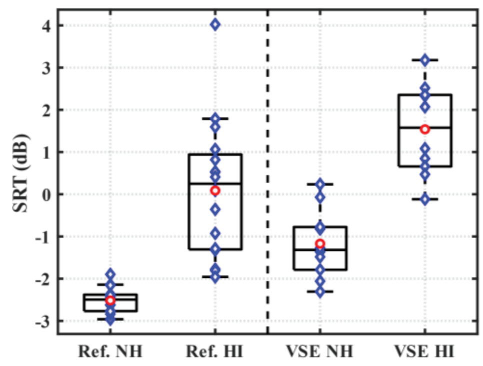


Fig. 7 – SRTs for reference and VSE

Sources	Mean squares	$p(F_{0.05})$
Hearing (NH – HI)	86.88	**<0.001**
Condition (Ref. – VSE)	24.06	**<0.001**
Hearing*Condition	0.03	0.87

Table 2 – Two-way ANOVA with interaction

A two-way analysis of variance (ANOVA) in Table 2 reveals significant differences between the means of each combination of hearing status (NH - HI) and condition (Ref. - VSE). However, there is no significant interaction effect between hearing status and condition. This indicates that the difference in performance between the NH and HI subjects is very similar in both conditions, at 2.74dB (Ref.) and 2.90dB (VSE). Likewise, the difference in performance between the reference and VSE conditions is very similar for both the NH and HI groups, at 1.17dB (NH) and 1.33dB (HI).

4. DISCUSSION

The results show that the difficulty of understanding speech in the HINT SI task is increased significantly by conducting it inside an auralized CSS, as SRTs were found to increase by over 1dB compared to the classical version, while retaining a stable difference between NH and HI SI performance. The wider interquartile range for NH subjects in the VSE condition can be explained by the more variable levels occurring in the office background noise. The fact that the interquartile range for the HI is largest in the reference condition is most likely due to those subjects' more variable hearing thresholds.

The realistic SI task as defined here differs in two major ways from existing SI procedures. First, the background noise was constructed from a real-world in-situ recording instead of speech-shaped stationary noise. Second, target and background stimuli were spatialized, not presented diotically. Both modifications were intended to increasing experimental realism and ecological validity.

Considering the spectral characteristics of the stimuli, the modulation spectrum of the realistic background noise is similarly shaped to that of the HINT sentences, as opposed to the classical HINT noise. The presence of these modulations may lead to poorer perception of the speech stimulus in the realistic background, due to effects of modulation masking. On the other hand, using modulated maskers in an SI task can introduce "listening-in-the-dips" (17), increasing speech intelligibility. Therefore, the specific structure and contextual properties of the speech components in the noise must play a role as well, for instance through informational masking (18).

Besides the stimuli's spectral properties, their spatialization could also decrease SI performance. The diotic speech stimuli of the classical HINT do not contain any reverberation, a property which is widely known to decrease speech intelligibility (19). The spectral characteristics of the single-channel and auralized versions of the HINT sentences are very similar, implying that the spatialization preserves the long-term average and modulation spectral properties.

A more cognitively inspired argument in support of increasing realism in SI paradigms builds on the concept of experimental realism, as opposed to mundane realism, as defined in psychology. Mundane realism refers to experimental conditions that mimic those of the real world as closely as possible, whereas experimental realism indicates the extent to which subjects actually experience their conditions as realistic (20). Therefore, to obtain meaningful results from a subject, his or her perception of realism can be just as important as its objective realization.

Even though the VSE did not include visuals, based on a 5-point modified Likert questionnaire given to all subjects, the reproduced environment was found to be very realistic (4), and subjective speech understanding was very to extremely difficult for hearing-impaired subjects (4-5). Respective reasons for these ratings were familiarity with similar situations in daily life, and the acoustic and cognitive interference of the intelligible background speech with the stimuli that needed to be attended.

However, is not clear if this increased perceptual difficulty corresponds to the actually experienced ecological struggle of HI people. One way to relate SRTs to reported real-world SNRs is by looking at which percentage of these SNRs are below a specific SRT. This provides a quantitative measure of the proportion of real-world sound scenarios in which subjects' SI would be below 50%, assuming the SI task is representative of the real world.

Applying this strategy here shows that the mean SRTs for the VSE background intersect a cumulative Gaussian distribution calculated from SNR distribution data collected by Pearsons et al. and Smeds et al. (21) at respectively 11.8% and 9.3% for NH subjects. This means that a predicted intelligibility of less than 50% would only occur in about one out of ten real-world scenarios. For the HI, these values are 24.0% and 14.9%. Conducting the same procedure for the reference condition gives values of 6.8/6.1% (NH) and 10.7%/8.6% (HI), respectively. It is unclear what % correct in terms of speech intelligibility is necessary for communication in the real world, but if 50% is taken as a measure, these results indicate that the number of situations HI people struggle can be as high as one in four. Applying a threshold of 75% SI for proper communication raises these values by about another 5%. For SRTs obtained in an SI task to meaningfully quantify these struggles, this comparison can help determine accurate and representative paradigms. The elevated percentages that result from a more realistic SI task seem to be more in line with the amount of situational difficulties HI people report.

The VSE condition is limited in its acoustical accuracy by the recording and reproduction methods. However, the application of dual-band decoding, as used here, significantly increases the spatial aliasing frequency of an HOA system (22), and as said, the subjective perceptual impression of the reproduction was deemed highly realistic. As said, no visual or situational context was available to the subjects during the SI task, the presence of which should decrease SRT scores. Thus, an SI task that would include visuals might need to be made even more challenging acoustically to remain representative of real-world difficulties. As such, it is likely that the SI paradigm with a realistic background developed here could still be improved to yield higher SRTs, akin to reality.

5. CONCLUSION

In this study, a novel implementation of a speech intelligibility task was realized, driven by principles of ecological validity and experimental realism at every stage in its development.

It was shown that both NH and HI SRTs obtained in an HOA-reproduced office meeting CSS are, on average, over 1dB higher compared to the classical headphone-based HINT reference condition. A comparison of the used signals suggests that this increase in difficulty is primarily caused by the speech-like envelope modulations and intelligibility of the SI masking noise, as well as the reverberant properties of the spatialized stimuli. By quantitatively relating SRTs to cumulative distributions of real-world recorded SNRs, it followed that mild to moderate hearing impairments might lead to less than chance level speech intelligibility in over 20% of critical sound scenarios.

To conclude, the present work suggests that introducing controlled realism in a speech intelligibility paradigm can produce results that are more representative of real-world conditions. Such paradigms may be more appropriate when attempting to characterize hearing difficulties that are not experienced in a laboratory setting, but in the real world.

REFERENCES

1. Moore BC. Perceptual consequences of cochlear hearing loss and their implications for the design of tearing aids. Ear Hear 17(2); 1996. p. 133-161.
2. Dreschler WA, Verschuure H, Ludvigsen C, Westermann S. ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. Audiology 40(3); May-June 2001. p. 148-157.
3. Christensen et al. ODEON Room Acoustics Software User's Manual, Odeon A/S; July 2016.
4. Best V, Keidser G, Buchholz JM, Freeston K. An examination of speech reception thresholds measured in a simulated reverberant cafeteria environment. International Journal of Audiology, 54(10); 2015. p. 682-690.
5. Timmer BHB, Hickson L, Launer S. Ecological momentary assessment: Feasibility, construct validity, and future applications. American Journal of Audiology 26(3S); October 2017. p. 436-442
6. Wolters F, Smeds K, Schmidt, Christensen EK, Norup C. Common Sound Scenarios: A Context-Driven Categorization of Everyday Sound Environments for Application in Hearing-Device Research. The Journal of the American Academy of Audiology 27(7); July 2016. p. 527-540.
7. Wagener K, Josvassen JL, Ardenkjær R. Design, optimization and evaluation of a Danish sentence test in noise. International Journal of Audiology 42(1); 2003. p. 10-17.
8. Nielsen JB, Dau T. The Danish hearing in noise test. International Journal of Audiology 50(3); March 2011. p. 202-208.
9. Soli SD, Wong LNL. Assessment of speech intelligibility in noise with the Hearing in Noise Test, International Journal of Audiology, 47(6); June 2008. p. 356-361.
10. Smeds K, Wolters F, Rung M. Estimation of signal-to-noise ratios in realistic sound scenarios. The Journal of the American Academy of Audiology 26(2); February 2015. p. 183-196.
11. Moreau S, Daniel J, Bertet S. 3D sound field recording with higher order Ambisonics - objective measurements and validation of spherical microphone. Audio Engineering Society, 120th Convention Spring Preprints 1; May 2006. p. 356-379.
12. Smeds K, Wolters F, Larsson J, Herrlin P, Dahlquist M. Ecological momentary assessments for evaluation of hearing-aid preference. The Journal of the Acoustical Society of America 143; April 2018.
13. Elko, GW. em32 Eigenmike microphone array release notes (v18.0). mh acoustics; June 2014.
14. Kinnunen T, Lib H. An overview of text-independent speaker recognition: From features to supervectors. Speech Communication, 52(1); 2010. p. 12-40.
15. Bisgaard M, Vlaming S, Dahlquist M. Standard Audiograms for the IEC 60118-15: Measurement Procedure. Trends in Amplification 14(2); 2010. p. 113-120.
16. Dreschler WA, Verschuure H, Ludvigsen C, Westermann S. ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. International Collegium for Rehabilitative Audiology 40(3); May-Jun 2001. p. 148-157.
17. Peters RW, Moore BC, Baer T. Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. The Journal of the Acoustical Society of America 103(1); January 1998. p. 577-587.
18. Westermann A, Buchholz JM. The influence of informational masking in reverberant, multi-talker environments. The Journal of the Acoustical Society of America 138(2); August 2015. p. 584-593.
19. Harris RW, Reitz ML. Effects of room reverberation and noise on speech discrimination by the elderly. Audiology 24(5); 1985. p. 319–324.
20. Aronson E, Carlsmith JM. Experimentation in social psychology. The handbook of social psychology 2(2); 1968. p. 1-79.
21. Wu Yu, Stangl E, Chipara O, Hasan, SS, Welhaven A, Oleson J. Characteristics of real-world signal-to-noise ratios and speech listening situations of older adults with mild to moderate hearing loss. Ear and Hearing 39(2); March-April 2018. p. 293-304.
22. Oreinos C, Buchholz JM. Objective analysis of Ambisonics for hearing aid applications: Effect of listener's head, room reverberation, and directional microphones. Journal of the Acoustical Society of America 137(6); June 2015. p. 3447-3465.