Effects of rain noise from metal roofs on speech identification

Yan Chen\textsuperscript{1a}, Xin Li\textsuperscript{1b}  
\textsuperscript{1}Institute of public security, Jiangsu Academy of Safety Science and Technology, Nanjing, China  
\textsuperscript{2}Nari research institute, Nari Technology co., ltd, Nanjing, China  
\textsuperscript{a}chenyan163mail@163.com, \textsuperscript{b}lixin11@sgepri.sgcc.com.cn

Abstract—Recently metal roofs were widely utilized in rainwater harvesting, and noise produced by rain on metal roofs cannot be ignored. The goal of this article is to demonstrate the effects of rain noise from the metal roofs on speech identification. Two recordings of rain noise from metal roofs and a speech corpus simulated two actual acoustical listening environments. Two experiments were carried out to investigate the interference of rain noise on speech identification in two different acoustical listening environments. Two different simplified ordinary classroom model were carried out in the article. The experiment 1 was in a classroom with a metal roof and experiment 2 was in two point in a simplified ordinary classroom which is close to other building with metal roofs. Some results were shown speech identification all dropped significantly with SNR reduces.

1. Introduction  
Metal roofs were extensive used in the process of rainwater harvesting which is an important way of optimizing the usage of water resources and promoting sustainable development. Noise caused by rain falling onto metal roofs is a commonly occurring problem which attracted a lot of attention of scientists \cite{1–3}. The level of noise produced by raindrop impacting on metal roofs can exceed 70 dB \cite{4}. Most buildings with metal roofs generated a high level of noise in the interior spaces once raining. The negative effects of rain noise amplified by metal roofs were affirmed in people’s lives \cite{5-6}. People were aware about these issues and classified sound from rain as noise that produces negatives effects on their work or study. Almost 82 percent of the respondents agreed that they also need to speak loud or even shouting if they were to continue talking. However, researchers haven't quantified the degree of shouting.

Hence, the objective of this paper is to examine the effects of rain noise from the metal roofs on speech identification. The effects are analyzed based on listening experiment.

2. Preparation Experiment  
2.1. Listeners  
Ten normal hearing adult volunteers, including 5 females and 7 males, which were between 27 and 36 years old, participated in all experiments. None of them were familiar with the purpose of the test and the sentences used during the test. Before the test all listeners were given basic instruction.

2.2. Target phrases  
Target phrases in preparation experiment were taken from the coordinate response measure (CRM)
corpus [7]. The CRM corpus has been used extensively in previous studies in speech identification [8–10]. The corpus contains sentences spoken by eight different talkers: four males and four females. Each sentence in the corpus has the structure: “Ready ‘callsign’ go to ‘color’ ‘number’ now.” The corpus contains all possible combinations of eight callsigns (“arrow,” “baron,” “charlie,” “eagle,” “hopper,” “laker,” “ringo,” “tiger”), four colors (“blue,” “green,” “red,” “white”), and the numbers between one and eight. On each trial, the target sentence varies randomly from the CRM corpus. The target phrase varied from trial to trial. A set of test contains 50 trials and preparation experiment contains 4 sets per listener. All sets were generated by using MATLAB (Math Works Inc.).

2.3. Procedure and Result
In preparation experiment, target signals were presented via headphone using HRTFs (head related transfer functions) [11] in order to simulate that the speech are in front of listeners in anechoic environment. All listeners finish all sets of tests alone.

Everyone has above 99% correct rate. Although all listeners are Chinese, they have received English training more than ten years. Hence, all listeners are familiar with these words and effect of vocabulary comprehension is negligible.

3. Experiment 1

3.1. Virtual Room
The virtual room used in experiment 1 was simulated using the image (ray-tracing) method [12–13]. The room was 10 m long, 6 m wide and 3 m high. The absorption coefficient of the internal ceiling was 0.05, and all other absorption coefficients of the room internal surfaces were set to the same value 0.2. This virtual room represents a simplified ordinary classroom with a metal roof.

3.2. Stimuli and Procedure
The recording of rain noise in experiment 1 was conducted in a temporary workshop with a metal roof. The temporary workshop with no walls is common on the construction site. A microphone connected to a preamplifier was placed in the center of temporary workshop and about 1.6 m above the concrete floor.

Formation way of target phrases in every trial of experiment 1 were the same as preparation experiment.

As shown in Figure 1 and Figure 2, target phrases and rain noise were simulated in front of listener and from right above listener head in a simplified ordinary classroom with a metal roof respectively, where listener was in the center of the room (x-coordinate of 0 m, y-coordinate of 0 m). Binaural room impulse responses can be obtained by convolving HRTFs of the particular directions with the room impulse response functions. Binaural stimuli were created by convolving the target signal and rain noise with respective binaural room impulse responses. On each trial, the target sentence varies randomly from the CRM corpus.

The target phrase varied from trial to trial while the rain noise remained the same. A set of test contains 50 trials and experiment contains 4 sets at different signal-to-noise ratios (SNRs). All sets were generated by using MATLAB and Cool Edit Pro.

3.3. Results and Discussion
Table I charts the correct rates in five different SNRs in experiment 1. Comparing the results with the preparation experiment, rain noise significantly affects listeners' ability of obtaining information from speech. It is obviously that the performance of speech identification drops significantly while SNR reduces. Rain noise depends on the rain impact rate and the diameter and impact velocity of the rain drops. Especially in storm which brings strong rainfall intensity in short duration, rain noise may prevent people from communication, because people can’t obtain enough information from speech.
4. Experiment 2

4.1. Virtual Room
The virtual room used in experiment 2 was also simulated using the image (ray-tracing) method. The room has same size as experiment 2. But all absorption coefficients of the room internal surfaces were set to the same value 0.2. This virtual room represents a simplified ordinary classroom with a concrete roof.

4.2. Stimuli and Procedure
The recording of rain noise in experiment 2 was conducted between a temporary workshop with a metal
roof and a building. The temporary workshop was about 5 m away from the building. A microphone connected to a preamplifier was about 1 m from the building and about 1.2 m from the concrete floor.

Formation way of target phrases in experiment 2 was the same as preparation experiment.

As shown in Figure 3, there were two locations of listeners, in the center of the room (x-coordinate of 0 m, y-coordinate of 0 m) and near the window (x-coordinate of 2.7 m, y-coordinate of 0 m). Targets were positioned at the front of the classroom (x-coordinate of 0 m, y-coordinate of 4.5 m). Rain noise was from the window which is in the center of a side wall in a simplified ordinary classroom (x-coordinate of 3 m, y-coordinate of 0 m). Binaural room impulse responses and binaural stimuli were obtained by the same way. This virtual listening environment represents a typical case that the listening room has no metal roofs and other building with metal roofs is nearby.

![Fig.3 Location planform in experiment 2](image)

### 4.3. Results and Discussion

It can be seen from table II and table III that in such a condition that the listening room has no metal roofs and metal roofs of other building are nearby, rain noise also significantly affects listeners’ ability of obtaining information from speech. Similarly, it can be seen that the performance of speech identification drops significantly while SNR reduces. Correct rates are significantly better than those of experiment 1. The intrinsic mechanism for this phenomenon is that the ability of a listener extracting the content of the target in background noise improves when the target is separated horizontally from the noise. The correct rates of two locations are close to each other at SNRs of 4 dB, 0 dB, −4 dB, and −8 dB. While SNR is −12 dB, the correct rate of location 1 is lower than the correct rate of location 2. The reason for the phenomena was conjectured to both head shadow effect and binaural interaction effect. Reflections on the wall may also have influences on the phenomena.

| SNR  | Correct rate |
|------|--------------|
| 4 dB | 97.2%        |
| 0 dB | 92.7%        |
| −4 dB| 87.7%        |
| −8 dB| 82.1%        |
| −12 dB| 56.7%      |

Tab.2 Correct rates of five different snrs in locations 1 (0m, 0m) in experiment 2
Tab.3 Correct rates of five different snrs in locations 2 (2.7m, 0m) in experiment 2

| SNR   | Correct rate |
|-------|--------------|
| 4 dB  | 96.3%        |
| 0 dB  | 81.7%        |
| −4 dB | 71.6%        |
| −8 dB | 51.1%        |
| −12 dB| 30.4%        |

5. Conclusion
The results obtained in the present study explore the effect of rain noise from metal roofs on speech identification. Listeners in study have high level of English, so they are familiar with the phrases in all trails and effect of English vocabulary comprehension is negligible.

From the results, speech identification drops significantly while SNR reduces in a simplified ordinary classroom with a metal roof. Rain noise may prevent people from communication when it is raining heavily, because people can’t obtain enough information from speech.

In a simplified ordinary classroom which is close to other building with metal roofs, the performance of speech identification drops significantly while SNR reduces similarly. Correct rates are significantly better than those of experiment 1. The correct rates of two locations are close to each other at SNRs of 4 dB, 0 dB, −4 dB, and −8 dB. While SNR is −12 dB, the correct rate of location 1 is lower than the correct rate of location 2. The intrinsic mechanism for this phenomenon is that the ability of a listener extracting the content of the target in background noise improves when the target is separated horizontally from the noise. Both head shadow effect and binaural interaction effect may have influences on the phenomena.

Future work is to explore the influencing of different locations in classroom of experiment 2 and the intrinsic mechanisms for these phenomena. Rain intensity, water drop size and velocity, roof construction, and interior acoustical characteristics, which play important roles in these phenomena, should be studied in detailed.

Acknowledgments
This work was financially supported by the NSFC (National Nature Science Foundation of China) (No.51808266).

References
[1] Dubout P. (1969) The sound of rain on a steel roof, J. Sound Vib., vol. 10, 144–150.
[2] McLoughlin J., Saunders D. J. (1994) Ford R. D. Noise generated by simulated rainfall on profiled steel roof structures, Appl. Acoust., 42: 239–255.
[3] Coffeen R. C. (2007) A Comparison of interior noise levels produced by rain impinging on several commercial roof constructions, J. Acoust. Soc. Am., 122: 2991–2991.
[4] Mariam S., Faris A. (2005) Sound Impact Measurement on Various Types of Roof Systems, Proceeding of International Conference on Applied Physics and its Environmental Aspects, Solo, Indonesia.
[5] Idris M. F. M., Hamzah A. I. N., Ayob S. M. (2012) Occupants’ responses on rain noise underneath metal deck roof system, Procedia Soc. Behav. Sci., 35: 404 – 411.
[6] Idris M. F. M., Musa M. M., Ayob S. M. (2012) Noise Generated by Raindrop on Metal Deck Roof Profiles: It's Effect towards People Activities, Procedia Soc. Behav. Sci., 36: 485 – 492.
[7] Bolia R. S., Nelson W. T., Ericson M. A., et al, (2000) A speech corpus for multitalker communications research, J. Acoust. Soc. Am., 107: 1065–1066.
[8] Arbogast T. L., Mason C. R., Kidd G. J. (2002) The effect of spatial separation on informational and energetic masking of speech, J. Acoust. Soc. Am., 112: 2086–2098.
[9] Ihlefeld A., Shinn-Cunningham B. (2008) Spatial release from energetic and informational masking.
in a divided speech identification task, J. Acoust. Soc. Am., 123: 4380–4392.

[10] Marrone N., Mason C. R., Kidd J. (2008) Tuning in the spatial dimension: Evidence from a masked speech identification task, J. Acoust. Soc. Am., 124: 1146–1158.

[11] Brungart D. S., Rabinowitz W. R. (1999) Auditory localization of nearby sources. Head-related transfer functions, J. Acoust. Soc. Am., 106: 1465–1479.

[12] Allen J., Berkley D. (1979) Image method for efficiently simulating small-room acoustics, J. Acoust. Soc. Am., 65: 943–950.

[13] Peterson P. (1986) Simulating the response of multiple microphones to a single acoustic source in a reverberant room, J. Acoust. Soc. Am., 80: 1527–1529.