Localization in Real and Virtual Rooms

M. Kathleen Pichora-Fuller[1], Bruce A. Schneider[2], Murray Hodgson[3], Waqar-un-nissa Valiani[3]

[1]School of Audiology and Speech Sciences, University of British Columbia, 5804 Fairview Ave., Vancouver, BC, V6T 1Z3
[2]Department of Psychology, Erindale College, University of Toronto, 3359 Mississauga Rd., Mississauga, ON L5L 1C6
[3]Occupational Hygiene Program and Department of Mechanical Engineering
University of British Columbia, 3rd Floor, 2206 East Mall, Vancouver, BC V6T 1Z3

Introduction

Elderly listeners experience more difficulty understanding spoken language in non-ideal listening conditions than do younger adults. Specifically, they have more trouble in many everyday listening situations where background noise or reverberation conditions are unfavourable. Despite their difficulty understanding what is said in non-ideal conditions, elderly listeners often perform like younger adults in ideal listening conditions, such as when they are talking to one familiar person in a small, quiet room. Furthermore, many do not have clinically significant elevations in pure-tone thresholds, and existing clinical tests conducted in the artificial conditions of soundbooths are not useful in predicting an individual's performance in real-world communication situations. Therefore, it is important to devise new methods to allow researchers and clinicians to better evaluate how listeners perform in non-ideal, real-world conditions.

Testing listeners in actual acoustic conditions would be the most ecologically valid approach; however, precise control of the test stimuli would be jeopardized. Auralization[1], or the simulation of acoustic environments, is another approach that permits more realistic conditions to be created while maintaining precise control of test stimuli. As a first step in adopting the latter approach, we tested and compared the abilities of listeners to localize speech signals in a real room and in simulations of the same room.

Method

The Real Room: The real room was a 12.92' wide x 17.42' long x 8.83' high room, with one door and no windows, located in a modern building with research offices and labs. Eight loudspeakers were arranged in a circle, 45° apart, such that each loudspeaker was 5' from the listener seated in the center of the room, and at the same height as the head of the listener. The reflective characteristics of the surfaces (walls, floor and ceiling) and the reverberation times per octave band for the room were measured [2].

The Virtual Rooms: Four simulations of the real room were constructed using Tucker-Davis Technologies (TDT) hardware and modified software [2]. The first simulation included only the direct wavefront. The second included the direct wavefront and the first reflections from each of the six surfaces, where the surfaces were modelled using the frequency-specific reflective characteristics measured in the real room. The third included the direct wavefront and a reverberant tail, which was modelled using the average of the frequency-specific reverberation times that were measured in the real room. The fourth included the direct wavefront, the first reflections, and the reverberant tail. The four simulations were convolved with each of three head-related transfer functions (HRTFs) provided with the TDT Soundstage software. Thus, a total of 12 conditions were constructed (4 room simulations x 3 HRTFs).

Stimuli: A set of 20 soundfiles were used. In each soundfile was a 4-second segment of 8-talker babble. The segments were shaped with a rise-fall time of 100 msec. The RMS of the soundfiles ranged from 2.013 to 2.014 Volts. In the real room, soundfiles were played out of the loudspeakers at a sound pressure level of 70 dB SPL. In the virtual rooms, the software was used to assign one receiver position corresponding to the position of the listener in the real room, and 8 source positions corresponding to the 8 positions of the loudspeakers in the real room.

Subjects: Twenty-four normal-hearing, young adult listeners were paid for their participation in the study.

Procedure: Each of the 20 babble files was played out once from each of the 8 loudspeaker positions. The order of presentation of files at each position was random. On each trial, the subject pushed one of 8 buttons labelled in compass directions (N, NE, E, SE, S, SW, W, NW) to indicate where the sound source seemed to be located. In Experiment 1, subjects received no feedback. In Experiment 2, subjects received feedback indicating that the response was correct or, if it was incorrect, what the correct response should have been. Each subject completed both experiments in the real room in a single one-hour session. Similarly, 12 sessions were completed later in the virtual rooms, with one session for each of the 12 conditions, and with the order of conditions varied across subjects.

Results

In both the real and virtual rooms, without feedback, subjects were poorest at localizing sound sources behind them (S), with the most common error being a front-back (N-S) confusion. However, with feedback, performance improved in both the real and virtual conditions.

Discussion

Auralization is feasible for use in the laboratory. It provides useful information about the real-world abilities of listeners to localize sounds. Its usefulness for measuring other aspects of perceptual performance in research and clinical studies remain to be determined.

[1] M. Kleiner et al., "Auralization -- An Overview", *J. Audio Eng. Soc.* 41 861-875 (1993).

[2] W. Valiani et al., "Auralization of Speech-Communication Cues", *Can. Acoust.* (Summary, this issue).