Evaluating the effects of hearing protection on speech production in noisy environments

Douglas S. Brungart, Mary T. Cord, Nancy P. Solomon, Katie Dietrich-Burns, Kim Block

1 Audiology and Speech Center, Walter Reed National Military Medical Center, Bethesda, MD 02123
douglas.brungart@us.army.mil

Abstract

Many factors can influence the voice levels that talkers choose to use when they initiate conversations in noisy environments, including the desire to be understood clearly, the desire to communicate privately, and the distorted perception of the background noise and of their own voices that can occur when hearing protection devices are worn. In this study, we examined the impact that two different kinds of earplugs and four levels of room noise had on the voice levels of talkers who were asked to privately communicate short verbal messages to a nearby acoustic manikin. Frozen noise samples were used for the room noise, which allowed the speech samples recorded by the manikin to be extracted from the noise by direct waveform subtraction. The results show that talkers wearing earplugs consistently use lower voice levels in noise than they do when not wearing earplugs, even when the earplugs produce relatively little attenuation. This highlights the important impact that earplug-related distortions in the perception of one's own voice (which are commonly referred to as the occlusion effect) can have on the effectiveness of speech communication in noise.

Index Terms: shouted speech, hearing protection, occlusion effect

1. Introduction

Many variables can influence the vocal effort levels that talkers choose to deploy when they are communicating in noisy environments. Usually, the primary goal is to speak loudly enough to be heard by the person to whom they are speaking. However, in most cases, the talker does not want to speak any louder than necessary, both because it requires additional effort to speak loudly and because the talker may be concerned that other people in the environment might overhear the conversation and/or be annoyed by it. In military environments, soldiers may also be concerned about being detected by an adversary if they shout too loudly. Thus, in most cases, talkers in noisy environments are motivated by an underlying goal to speak just loudly enough to be clearly understood by their target audience.

The actual selection of this appropriate voice level is not trivial, even in the best circumstances, because it requires the talker to accurately estimate both the loudness of his or her own voice at the location of the listener and the level of noise at that location. However, this selection process is complicated even further in cases where the person talking is required to use some type of hearing protection. Hearing protection has two important effects on how the environment is perceived by the talker. The first and most obvious impact is that it attenuates the perceived level of noise in the room. However, it also has the more subtle effect of distorting the way the talker perceives his or her own voice. Almost everyone who has used earplugs knows that they make one's own voice sound “boomy” or “hollow”, similar to what you might expect to perceive if you were talking with your head inside a barrel. This phenomenon, which is known as the “occlusion effect”, is caused by the transmission of the talker’s voice through bone conduction into the constricted ear canal space confined by the earplug. Acoustically, its effects can be estimated by placing a microphone in the ear canal and comparing the speech levels measured for a talker speaking at the same level with or without the earplug inserted. Measurements of this type show that the insertion of an earplug can increase the loudness of a talker’s voice at low frequencies (250-500 Hz) by 20 dB or more [1]. Occlusion can also be measured subjectively, by obtaining the talkers rating of the hollowness of his or her own voice. Alternatively, bone conduction thresholds can be used to determine the minimal detectable intensity of a 500 Hz tone generated by a bone-conduction transducer with and without the earplug inserted (in which case the insertion of an earplug will produce a 10-20 dB decrease in the minimum detectable threshold).

From these results, one might expect the use of hearing protection to have two consequences for speech production in noise: 1) to reduce the perceived level of noise in the surrounding environment; and 2) to increase the talkers self-perceived vocal loudness for a given amount of vocal effort. Not surprisingly, the net effect of these distortions is a general tendency for talkers using hearing protection in noise to speak more quietly then they normally would if they were speaking in the same environment with open ears. This was demonstrated by Tufts and Frank [2], who asked talkers wearing foam or flange earplugs to produce speech in various levels of simulated noise. However, in order to obtain clean recordings of the talker’s voice that were not contaminated by the room noise, that study was conducted in a quiet room with noise-producing headphones worn over the earplugs, rather than in an actual noise-filled room. This made it impossible to evaluate the impact of earmuffs on speech production in noise, and it meant that the “open ear” condition was not truly an “open ear in noise” condition but rather a “noisy headphone in quiet” condition. In this study, we conducted a similar experiment using reproducible frozen noise to adjust the level of a diffuse noise field in an audio test booth. This allowed us to use waveform subtraction to recover the noise spectrum of the talker’s voice even in cases where the subject was truly speaking with unoccluded ears in a noisy environment.

2. Hearing Protection Devices

The two earplug conditions tested in the experiment were the two possible configurations of the EAR Combat Arms Earplug (CAE). The CAE is a passive nonlinear earplug system that is designed to protect listeners from high-level impulse noises...
The configuration of the testing room is shown in Figure 3. In the Open configuration, the CAE earplugs produce the same amount of occlusion but substantially less attenuation than the CAE Closed condition at that frequency (32 dB vs 10 dB). Thus, it appears that the CAE Open earplug produces the same amount of occlusion but substantially less attenuation than the CAE Closed condition.

### 3. Speech Production Experiment

Once the occlusion effect for each earplug condition was determined, an experiment was conducted to determine the impact that the CAE earplugs had on speech production in noise. Talkers in a noise-filled sound booth were asked to privately communicate short verbal messages to a nearby acoustic manikin. The configuration of the testing room is shown in Figure 3. The
room itself was a standard sound booth, 2.75 by 2.6 m by 2.3 m, with a total of 16 loudspeakers (one centered on each wall and in each corner at ear level, and one in each upper and lower corner). For the purposes of this experiment, the talker (subject) was seated on a stool that was 1 m away from the location of a Knowles Acoustics Manikin for Acoustic Research (KEMAR) who was placed at the same height as the subject but was rotated away so the talker would be speaking at a point roughly 45° behind the manikin’s right ear. This placement was used to replicate a situation where the talker had an urgent message that needed to be communicated verbally because the listener was facing in another direction and therefore could not be alerted to the message through hand gestures or other non-verbal means.

In order to further constrain the talker’s level of vocal effort, a mirror was placed behind the KEMAR in such a way that it reflected an image of a second manikin that appeared to be located 2.4 m away from the talker. This image was identified to the subjects as an eavesdropping listener, and they were instructed that their messages should be spoken loudly enough to be understood by the nearby target manikin, but not by the more distant eavesdropping manikin.

The experiment was divided into six 20-trial blocks, with each subject participating in two blocks with each type of hearing protection (open ear, CAE Open, or CAE Closed). At the start of each block, the subject was instructed to produce a 5-s quiet speech sample by reading a short passage about hurricanes. This speech sample was recorded both by the KEMAR in-ear microphones and by a reference microphone (Larson Davis 2527) located 0.5 m in front of the listener’s mouth. Then, prior to the first trial of the experiment, the experimenter provided the subject with a number from 1 to 20 that corresponded to one of 20 quasi-militarily-relevant test phrases (e.g., “Sgt. Ke-mar, we have a 10-91 in the Green Zone”) printed on posters that were mounted on the wall of the sound booth behind the KEMAR manikin. Then the subject sat silently while a 4.5 s sample of noise was generated by all 16 loudspeakers in the room and recorded by the in-ear microphones. This noise sample was a diffuse noise filtered to produce a pink noise spectrum at one of four different sound levels (50, 60, 70 or 80 dB SPL) at the location of the reference microphone. After the completion of this noise interval, there was a short pause, and then the exact same 4.5 s frozen noise sample was played through the loudspeakers in the room a second time. During this second interval, the subject read the designated target phase at a level that would be understandable by the nearby manikin but difficult to understand for the more distant eavesdropping manikin. This second speech-plus-noise sample was also recorded; the initial noise-only recordings were digitally subtracted from the speech-plus-noise recordings to recover a noise-free estimate of the speech signal at the locations of the three microphones in the room. Then the subjects were asked to give a verbal estimate of the amount of vocal effort used to produce the speech on a numerical scale ranging from 1 (not at all effortful) to 10 (extremely effortful).

The extracted speech waveforms were played and displayed to the experimenter, who used a mouse to manually extract the target phrase from the full 4.5 s recording. Recordings that were invalid (for example, because the subject made noise during the initial reference noise recording) were discarded, and randomly re-collected at some point before the end of the block. This procedure was repeated until a total of 5 speech samples were collected at each of the four noise levels.

Figure 4 displays examples of the magnitude spectra of the speech signals recorded during the experiment. The top row shows a case where the noise was set to 80 dBA, and the bottom row shows a case where the noise was set to 50 dBA. In both cases, the left panel displays the recording of the noise made in the first interval of each trial when the subject was sitting quietly, and the middle panel shows the recording in the second interval when the subject was asked to speak over the noise. The right panel replots these two curves, along with a third curve (solid black line) showing the difference between the two samples. This effectively represents a clean representation of the target speech signal. The gray line is the noise floor in each condition, calculated from the difference between two frozen noise samples when no speech signal was present in the second interval. Note that, in the 80 dB case, the technique is capable of recovering speech signals well below the level of the noise, and that, even at high frequencies (8 kHz), the speech recordings were well above the inherent noise floor of the waveform subtraction technique used to collect the speech samples.

4. Results

Figure 5 illustrates the relationship between the overall level of the speech samples produced by the subjects (as measured by the reference microphone 0.5 m away from the listener) and the overall level of the noise in the room. A few things are notable in this figure. First, it is apparent that, even in the open ear condition, the subjects generally did not increase their vocal intensity enough to maintain a constant signal-to-noise ratio in the high-noise conditions of the experiment. Second, it is apparent that both hearing protection conditions caused the listeners to reduce the intensity of their voices relative to the open-ear condition. In the 80 dB noise condition, subjects wearing the CAE Open earplugs spoke 5.8 dB less intensely than they did when they wore no hearing protection, and in the CAE Closed condition they spoke 10.1 dB less intensely than they did with no hearing protection. Notably, this 10 dB reduction in the CAE Closed is almost exactly the same as that reported by Tufts and Franks [2] for very similar flanged earplugs.

A more relevant estimate of the impact the hearing protectors had on speech communication in noise is shown in Figure 6, which shows the predicted speech intelligibility of the speech signals at the location of the manikin’s better ear (the one oriented toward the talker) as a function of noise level and hearing protection condition. These estimates were obtained with the...
The Listening Talker, Edinburgh, 2–3 May 2012
Oral session 2, May 3rd, 16.00

noise and speech measurements made from the in-ear microphone to calculate the Articulation Index (AI) in each condition [4], and then convert the AI score into an estimated percent intelligibility for NU6 words. In this figure, we can see that the subjects clearly did not increase their vocal intensity enough to maintain a high level of intelligibility when the noise level was above 60 dB, even in the open ear condition. The use of hearing protection in these high noise conditions had a devastating impact on predicted intelligibility in the 80 dB noise environment, with predicted NU6 scores dropping from 55% correct to 38% correct with the CAE Open earplugs, and plummeting to 18% with the CAE closed earplugs.

5. Conclusions

In this study, we examined the impact that hearing protection devices have on the production of speech in noisy environments. This study contributes to extant literature on this topic by specifically examining the impact of attenuation and occlusion for passive non-linear earplugs such as the CAE, which are often described during training as providing protection from high levels of impulse noise while being “acoustically transparent” in other low-level noise environments. Although these non-linear earplugs do produce relatively low attenuation, especially at low frequencies, the results of this experiment demonstrate that they actually produce just as much occlusion as a normal hearing protection device. When individuals wear these devices in environments that are noisy but not necessarily hazardous (i.e. < 85 dBA), it is highly probable that these earplugs will bias them to speak more quietly than necessary to achieve effective communication. Therefore, training may be needed to make talkers aware of this potential issue and ensure that they do not incorrectly attribute this effect to some problem with the attenuation characteristics of the device. These results also indicate that the measurement technique described here, which uses waveform subtraction to recover the speech waveform produced by a talker in a noisy environment, may have many advantages over other methods of examining the effect of background noise on speech production. This technique could have many potential applications in future studies in this important area.

6. Acknowledgments

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Army, or Air Force, the Department of Defense, nor the U.S. Government.

7. References