USAARL Report No. 2004-13

Speech Intelligibility in Noise Using Throat and Acoustic Microphones

By Barbara E. Acker-Mills, Adrianus J. M. Houtsma, and William A. Ahroon

Aircrew Protection Division

April 2004

USAARL

U.S. Army
Aeromedical Research
Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
April 2004

3. REPORT TYPE AND DATES COVERED
Final

4. TITLE AND SUBTITLE
Speech Intelligibility in Noise Using Throat and Acoustic Microphones

5. FUNDING NUMBERS
PR: 878
PE: 622787
TA: O
WU: DA360347

6. AUTHOR(S)
Barbara Acker-Mills, Adrianus J.M. Houtsma, and William A. Ahroon

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577
Fort Rucker, AL 36362-0577

8. PERFORMING ORGANIZATION REPORT NUMBER
2004-13

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
504 Scott Street
Fort Detrick, MD 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release, distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Rotary-wing aircraft crew are exposed to high levels of ambient noise that can cause hearing loss and impede speech communication. Existing technology generally focuses on the listener and devices that can increase the signal-to-noise ratio and decrease noise-induced hearing loss. However, speech intelligibility is still dependent on the quality of the original speech signal. Throat microphones greatly reduce ambient noise in the original signal, thus enhancing the signal-to-noise ratio for the listener, but higher frequencies are eliminated which may negatively affect speech intelligibility.

Speech intelligibility for signals generated by an acoustic microphone, a throat microphone, and the two microphones together was assessed using the Modified Rhyme Test (MRT). Stimulus words were digitally recorded in a reverberant chamber with broadband noise intensity at 90 dB(A) and 106 dB(A). Listeners completed the MRT task in the same settings, thus simulating typical environments of a rotary-wing aircraft. Results show that speech intelligibility is much worse for the throat microphone than for the acoustic microphone, particularly for the higher noise level. In addition, no benefit is gained by adding the throat microphone signal to the acoustic microphone signal.

14. SUBJECT TERMS
Noise, ambient noise, hearing loss, speech intelligibility, acoustic microphone, throat microphone

15. NUMBER OF PAGES
16

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
Unclassified

18. SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

19. SECURITY CLASSIFICATION OF ABSTRACT
Unclassified

20. LIMITATION OF ABSTRACT
Unlimited

Acknowledgements

We thank Mr. Christopher Madsen for his assistance supporting the software used to conduct this research, CPT Martin Robinette for his assistance in conducting this study, Ms. Janet Mauldin for her usual expert secretarial support, and our editor, Ms. Linda Burt. This research was funded in part by the U.S. Army Medical Research and Materiel Command In-house Laboratory Independent Review initiative.

Table of contents

Page

Notice

Qualified requesters

Change of address

Disposition

Disclaimer

Human use

Human subjects participated in these studies after giving their free and informed voluntary consent. Investigators adhered to AR 70-25 and USAMRMC Reg 70-25 on Use of Volunteers in Research.

Introduction

The ability to communicate effectively is of paramount importance in the rotary-wing aircraft environment. Effective communication leads to increased aircrew safety and performance resulting in successful mission completion. As communication is degraded, mission capability is reduced and the safety of the aircrew is compromised. A communication system involves at least one "talker" (sender), one "listener" (receiver), and any equipment used to augment or transmit information. Most research focuses on the listener and devices that can increase the signal-to-noise ratio and reduce noise-induced hearing loss.

The standard U.S. Army aviator helmet, the HGU-56/P Aircrew Integrated Helmet System, is designed primarily for impact protection. This helmet also includes sound-attenuating earcups, a noise-canceling microphone (the boom microphone), and an optional Communication Earplug (CEP)*. Thus, the helmet improves the signal-to-noise ratio (for the listener) and also protects the crew from noise-induced hearing loss (for the listener and talker).

Whereas a large corpus of information exists on the listener component of the communication system, very little research assesses problems at the talker level, and no device used with the HGU-56/P effectively enhances the talker's message under a wide range of conditions. The noise-canceling boom microphone in the HGU-56/P works well if positioned and used properly, but microphone use and noise conditions may impair performance. An open microphone, in contrast to the usual "keyed" microphone, may have to be used in some situations that require use of both hands (e.g., a crew chief operating a hoist may need to use both hands on the hoist control and cable). There are some aircraft environments in which an open microphone is the normal operating condition (e.g., the British Army). Noise conditions may be encountered (e.g., air movement from an open window or door) for which the normal noise-canceling microphone was not designed to minimize. In these conditions, a throat microphone could be used to improve intelligibility.

Aircrew during World War II faced similar noise issues, using more bulky, less sophisticated, acoustic microphones. Acoustic microphones were either hand-held or mounted on a chest plate. Signals were very noisy and these microphones were impractical for manual operations or for tasks requiring excessive head motion (Martin, 1947). Throat microphones were developed in response to these problems and used during the later years of WWII.

Since acoustic microphones convert sound energy (spoken words) into electrical energy to be transmitted into the communication system, they must necessarily also transmit any ambient noise around the talker's mouth. Removing this ambient noise from the communication signal should improve the signal-to-noise ratio for the listener. Throat microphones are designed to pick up (transduce) the vibrations of the vocal apparatus at the throat instead of the vibrations of air molecules at the mouth. Thus, if throat microphones are isolated from airborne sound (by

---

* See Appendix A, Manufacturer's list

contact with the skin) and/or are less sensitive to airborne sound than standard microphones, the speech signal input into the communication system should contain less noise.

Whereas the signal-to-noise ratio with throat microphones is enhanced, speech intelligibility may actually be worse with these devices. Consonant sounds are produced with the articulators (the tongue, lips, jaw position, etc.) and, because information is picked up before the level of the articulators, throat microphones should not be very effective in transmitting consonant sounds. However, ambient noise is not transmitted and vowel sounds (produced by the vocal cords) should be transmitted very effectively. The boom microphone, present at the level of the articulators, transmits consonant sounds, but with the trade-off of also transmitting ambient noise, resulting in a lower signal-to-noise ratio. It is possible, therefore, that speech intelligibility may be enhanced when both microphone types are used together; the boom microphone will add consonant sounds to the relatively noise-free, prominent vowels of the throat microphone.

During the development of throat and other contact-type microphones, there was little formal, systematic evaluation of speech intelligibility in noise using the devices, and the few existing studies contain conflicting results. Snidecor, Rehman, and Washburn (1959) explored vowel intelligibility with contact microphones located on different areas of the head and neck. Stimuli were recorded in quiet and presented over headphones to listeners who also were in a quiet environment. The authors rated intelligibility and another group of listeners gave quality judgments of the vowels. Contact microphone locations at the forehead, mastoid, and larynx were highest in intelligibility and quality ratings. While somewhat informative, this study does not use objective speech intelligibility measures, and with all recording and listening completed in a quiet environment, does not assess how contact microphones might function in noise.

Oyer (1955) evaluated intelligibility of words recorded with an acoustic microphone and a microphone placed in the ear canal. The ear microphone essentially is a contact microphone, as it transmits vibrations in the skull created by speech sounds. Air traffic control words in carrier phrases were recorded in quiet and mixed with 74 dB white noise prior to being presented to subjects over standard headphones. The signal-to-noise ratio was manipulated by attenuating the speech signal (-12, -15, and -18 dB). Results revealed a microphone × signal-to-noise ratio interaction, where speech intelligibility decreased for both microphones as the signal-to-noise ratio decreased, but the decrement was less for the ear microphone. Although not part of the formal study, it was reported anecdotally that simultaneous presentation of the acoustic and ear microphone stimuli resulted in very good speech intelligibility.

A study by Moser and Dreher (1956) is most applicable to the current research project. They used a noise-canceling acoustic microphone, an ear microphone, and a bone conduction microphone placed on the forehead to record the Phonetically Balanced (PB) word lists [developed by Egan (1948) and still specified by ANSI S3.2-1989 (R1999), *American Standard Method for Measuring the Intelligibility of Speech Over Communication Systems* (1989)]. The words were recorded by pilots in two different transport aircraft. Ambient noise in a KC-97 aircraft was measured at 97 dB and was measured at 106 dB in a C-124 aircraft. Listeners were in a quiet environment and stimuli were presented a 77 dB over standard headphones. Listeners became familiar with the two different microphone stimuli by listening to a paragraph followed by operational instructions. They then completed the PB task. All microphone transmissions were judged as "acceptable" during the familiarization phase, but PB results were better with the acoustic microphone than either the bone or ear microphone in both aircraft environments. An informal evaluation of bone and ear receivers (not microphones) found ear receivers to be rated as "excellent" and the bone conduction receiver (placed on the mastoid) to be considered "not acceptable" if the ears were not shielded from the noise.

In view of the sparse experimental research concerning the original contact-type microphones, including throat microphones, it is surprising that very few studies address speech intelligibility using modern throat microphones or similar devices. In fact, a thorough literature search found only one peer-reviewed paper (Horie, 2002) that mentioned speech intelligibility secondary to a study of temporary threshold shifts (TTS). This study evaluated use of an active noise reduction system and a bone conduction microphone contained in the helmet of steel workers. Although TTS was reduced only at 4 kHz for this system compared to an earphone, earplug, and acoustic microphone system, workers subjectively rated the new system as “superior” in seven areas, including “clearness of voice.” No objective tests of speech intelligibility were completed. A few recent conference presentations have discussed bone-conduction communication systems (e.g., Letowski, 2003), but these studies have not yet been published, nor did the listening conditions approximate the noisy environment of rotary-wing aircraft.

Even in the absence of speech intelligibility data, throat microphones are marketed to law enforcement agencies, fire departments, and to a variety of users for applications that require special environmental controls (respirators, hazardous material suits, etc.) or extremely rugged construction (waterproof, dustproof). Elements of the Department of Defense are very interested in communication systems for noisy environments, but systematic research evaluating the intelligibility of speech transmitted with throat microphones must be completed before recommendations can be made for their use.

The current study provides an objective, experimental evaluation of speech intelligibility for stimuli recorded using the HGU-56/P acoustic microphone, a commercially available throat microphone, and the two microphones used together. The experimental conditions include realistic noise conditions and thus address the feasibility of use of the three microphone options in rotary-wing aircraft.

Materials and methods

Speech intelligibility was measured with the Modified Rhyme Test (MRT) using procedures specified by ANSI S3.2-1989 (R1999), *American Standard Method for Measuring the Intelligibility of Speech Over Communication Systems* (1989). This test utilizes 300 stimulus words arranged in 50 six-word ensembles, where words in individual ensembles differ only in initial or final consonant sound. A visual display of all six words is presented, followed by the spoken phrase "Select the word [target word] please," and one of the six words played over the communication system. The listener chooses which word was heard.

Stimuli

The 300 MRT stimuli ("Select the word [target word] please") were recorded in a reverberant chamber with background broad-band noise intensity at 90 dB(A) and 106 db(A). The higher level simulates a UH-60A Black Hawk helicopter in doors closed, straight-and-level flight at 120 knots indicated airpseed (KIAS). A single-transducer LASH II throat microphone (based on the Thales Acoustics* RA440 throat microphone) was used in conjunction with the HGU-56/P noise-canceling acoustic microphone to record the stimuli. This microphone is similar to the throat microphones used by the Navy SEALS and Army Special Forces. The LASH II throat microphone is a lightweight (55 gm excluding plug), small, rugged device with medium sensitivity (-47 dB re 1V/Pa) specifically designed for use in very high noise environments (e.g., 120 dB) such as rotary-wing aircraft. The design includes special attention to the requirements of speech intelligibility (frequency response 150 – 5000 Hz).

The male talker wore a throat microphone and an HGU-56/P with the standard noise-canceling acoustic microphone. The talker fastened the throat microphone at a comfortable position and pressure, which was measured at about 200 grams of force. (Thales Acoustics does not provide specific directions for use of the throat microphone, except that it should fit comfortably without undue pressure.) The sound-attenuating earcups of the HGU-56/P and the use of CEPs protected the talker from noise during the recording session. Each phrase could be no longer than two seconds. If this time limit was not met, the phrase was recorded again.

Two separate analog-to-digital channels (40k sampling rate) were used to record stimuli simultaneously from the two separate microphones. Two separate sound files were created for each stimulus; one from the boom microphone and one from the throat microphone. Stimuli were post-processed to ensure equivalent overall root mean square (RMS) levels within each microphone type. The stimulus presentation/data collection program combined the two sound files for the boom plus throat microphone conditions. Because the stimuli were recorded simultaneously, the timing parameters for the two stimuli were exactly the same, thus avoiding any timing problems when the signals were combined.

Several stimuli randomly were selected for spectrum analysis. Because each of the target words contained different types of consonants that differ in intensity during natural speech, the signal-to-noise ratios were not exactly the same. However, the signal-to-noise ratios for stimuli recorded from the throat microphone were always higher (approximately 10 dB) than for stimuli from the boom microphone. In addition, the talker consistently spoke louder when in the 106 dB(A) noise environment (the Lombard effect). Thus, for each microphone type, the signal-to-noise ratios were similar between the 90 dB(A) and 106 dB(A) stimuli. Sample waveforms and spectra can be found in Appendix B.

Participants

Participants were Soldiers at Fort Rucker, Alabama, awaiting the Army Warrant Officer Course and flight training school. Nine males (average age = 25) and one female (age 29) volunteered for the study. All but one volunteer had normal hearing as confirmed by recent physical exams or by audiograms performed at the U.S. Army Aeromedical Research Laboratory (USAARL) acoustics laboratory. One male (age 43) reported having tinnitus in his right ear. Each participant read and signed an informed consent form.

Procedure

A 3 × 2 × 2 repeated-measures full-factorial design was used, with microphone type (boom, throat, boom plus throat), recording noise [90 dB(A) and 106 dB(A)], and listening noise [90 dB(A) and 106 dB(A)] as the independent variables. One block contained 50 six-word visual displays, with microphone type, recording noise, and listening noise held constant within a block. The MRT auditory target words were randomized between the 12 blocks to avoid learning effects. Each block lasted approximately 4 minutes and 20 seconds, and running order of the specific conditions was randomized among the 10 participants.

All testing took place in the USAARL-Acoustics Laboratory reverberant chamber, and stimulus presentation/response collection was coordinated using Tucker-Davis System II* psychoacoustic modules equipment and custom-written software. The purpose of the study and the experimental conditions were explained to the participant, then the participant was fitted with an HGU-56/P. Earplugs were not used, as the earcups of the helmet provided sufficient hearing protection for exposure to the 106 dB(A) noise for the 52-minute duration of the study. The six-word MRT visual displays were presented on a computer monitor, followed by the phrase "Select the word [target word] please" presented in the earphones. Participants had five seconds in which to use a mouse to select the target word that was spoken. A "no response" was recorded if the trial timed out or if the "Skip" button was selected.

Results

The average MRT percent correct results are displayed in Figure 1 (error bars represent standard error of the mean in this and all subsequent figures). Conditions are labeled according to the following three-digit code: Digit 1 = microphone type, Digit 2 = recording noise, and Digit 3 = listening noise. The Table illustrates the coding used in Figure 1.

Table.
Key to code used in Figure 1.

Microphone type:	1 = boom	2 = throat	3 = both
Recording noise:	1 = 90 dB(A)	2 = 106 dB(A)	
Listening noise:	1 = 90 dB(A)	2 = 106 dB(A)	

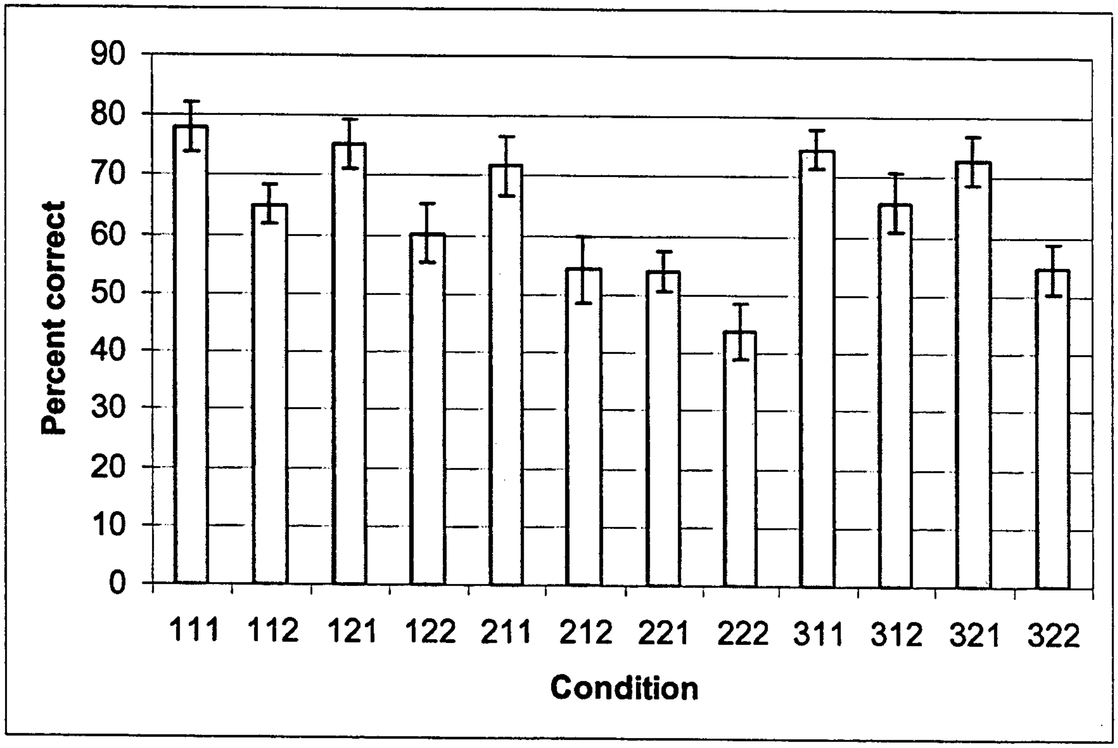


Figure 1. Average performance for individual MRT conditions.

These results were analyzed using a repeated-measures analysis of variance (ANOVA). The ANOVA table can be found in Appendix C, and individual subject data can be found in Appendix D. Subject 10 was the person who reported tinnitus in the right ear; while his percent correct scores were lower than the other subjects' scores, the patterns were the same, and excluding his results from the analysis did not change statistically significant findings. Main effects for all three independent variables were observed. Speech intelligibility was reduced significantly for stimuli recorded with the throat microphone, and intelligibility for the boom microphone alone was not significantly different from the boom and throat microphone combination (Figure 2).

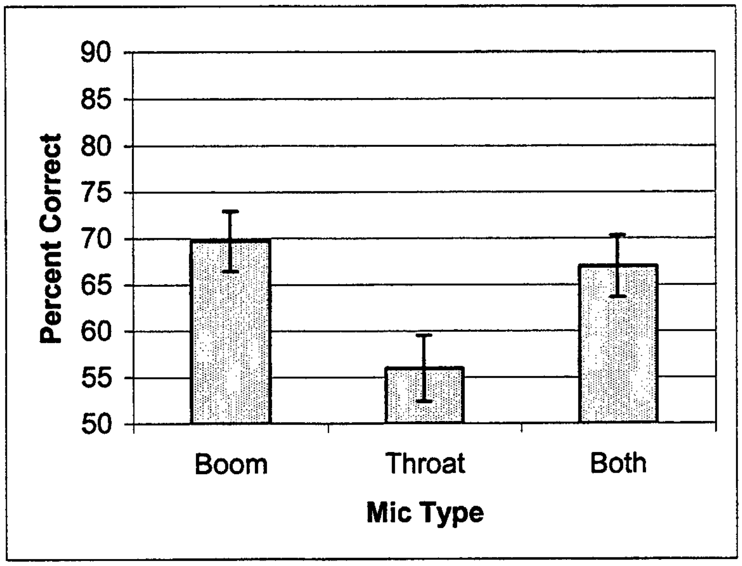


Figure 2. MRT performance as a function of microphone type.

As expected, performance was significantly worse for stimuli recorded in the 106 dB(A) noise compared to the 90 dB(A) noise and for the 106 dB(A) listening noise compared to the 90 dB(A) noise. Both of these main effects can be seen in Figure 3.

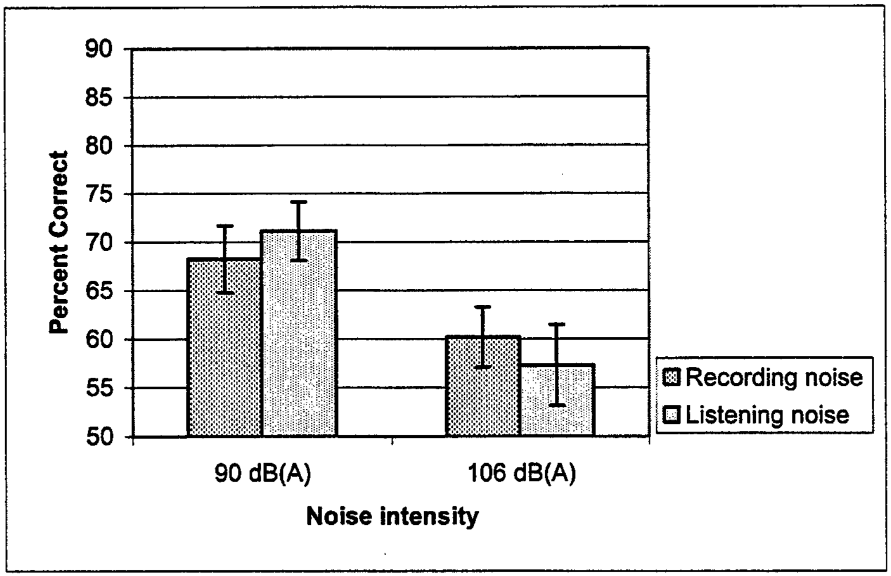


Figure 3. MRT performance as a function of noise intensity.

Finally, Figure 4 shows an interaction of microphone type and stimulus recording noise, where performance drops significantly for the throat microphone when recording noise increases from 90 dB(A) to 106 dB(A). Neither the boom microphone alone nor the boom and throat microphone combination experience this performance decrement.

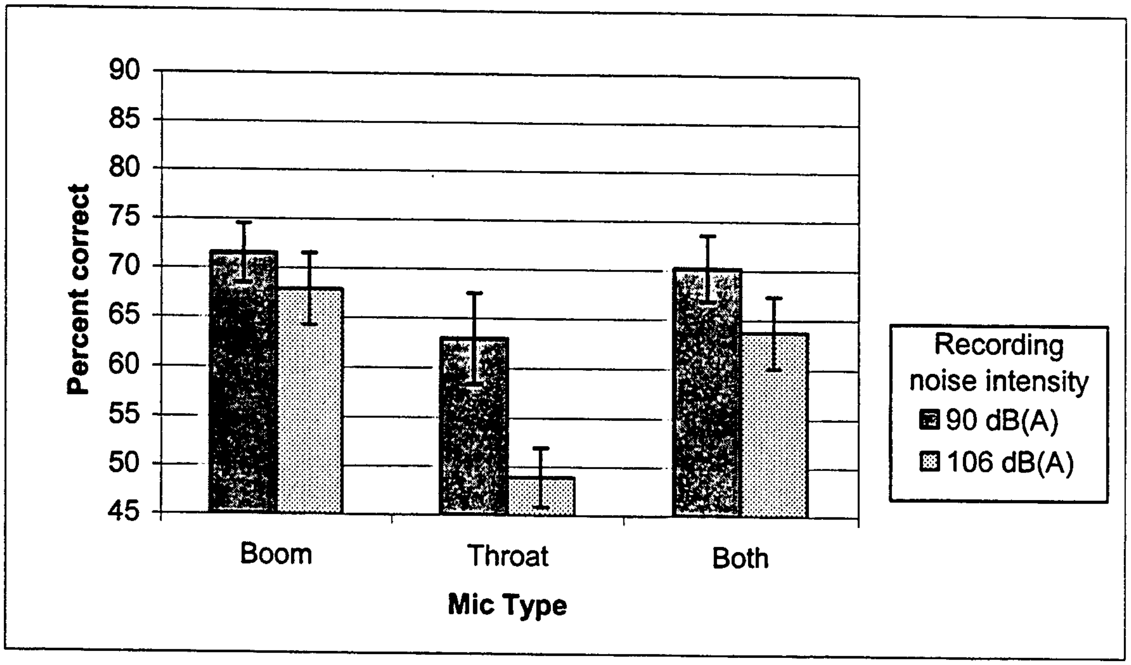


Figure 4. MRT performance as a function of microphone type and recording noise intensity.

Discussion

Research addressing the problem of noise and its detrimental effects on communication and hearing loss typically focuses on the "listener" (receiver). Whereas devices such as helmet earmuffs and the Communication Earplug can be useful (especially for hearing protection), speech intelligibility is still dependent on the quality of the original signal produced by the "talker" (sender). As noted above, the effectiveness of the noise-canceling boom microphone is reduced under various flying conditions that create unpredictable and highly variable noise. If this unpredictable ambient noise could be eliminated in the transmitted speech signal, the signal-to-noise ratio would be enhanced, and speech intelligibility also might be improved. Throat microphones greatly reduce or eliminate ambient noise because the microphone is in direct contact with the throat. Thus, the signal-to-noise ratio is better than that of an acoustic microphone used in a noisy environment.

The current results clearly demonstrate that while the throat microphone enhances the signal-to-noise ratio, the accompanying reduction in higher frequencies (see Frequency Spectra in Appendix B) degrades speech intelligibility compared to the acoustic boom microphone. It is presumed that the lack of consonant information in the speech signal is the cause of reduced speech intelligibility. These results are troubling in that the MRT task represents a closed set of words and should be a condition where intelligibility is best (Suter, 1989). Whereas normal flight procedures also have standard communication phrases (basically a closed set), non-standard speech will most likely occur in emergency/high panic situations. These are precisely the situations in which good speech intelligibility is critical, and the current results show that intelligibility using throat microphones is not adequate under the most benign of situations. In addition, throat microphones are most deficient in high levels of noise, precisely the environment for which they are designed. Finally, speech intelligibility is not enhanced when a throat microphone is used in conjunction with the boom microphone.

Conclusions

The current study demonstrates that the use of a throat microphone in noisy environments similar to that of rotary-wing aircraft does not increase speech intelligibility. Thus, it is recommended that at this time, the U.S. Army not consider the use of throat microphones in noisy environments. It is possible that future technology will improve throat microphone performance, but consonant information will never be able to be transmitted adequately if speech information is picked up only from the throat area.

References

American National Standards Institute. 1989. Method for measuring the intelligibility of speech over communication systems. Melville, NY: Standards Secretariat, Acoustical Society of America. ANSI S3.2-1989 (R1999).

Egan, J.P. 1948. Articulation testing methods. Laryngoscope. 58(9): 955 – 991.

Horie, S. 2002. Improvement of occupational noise-induced temporary threshold shift by active noise control earmuff and bone conduction microphone. Journal of Occupational Health. 44(6): 414 – 420.

Letowski, T. 2003. Bone conduction communication: Applications and limitations. Paper presented at the National Hearing Conservation Association 28th Annual Conference, February 20-22, at Dallas, TX.

Martin, D.W. 1947. Magnetic throat microphones of high sensitivity. Journal of the Acoustical Society of America. 19(1): 43 – 40.

Moser, H.M. and Dreher, J.J. 1956. Operational tests of miniature microphones and receivers. Columbus, OH: The Ohio State University Research Foundation. Technical Report No. 36 for Contract No. AF 19(604)-1577.

Oyer, H.J. 1955. Relative intelligibility of speech recorded simultaneously at the ear and mouth. Journal of the Acoustical Society of America. 27(6): 1207 – 1212.

Snidecor, J.C., Rehman, I., and Washburn, D.D. 1959. Speech pickup by contact microphone at head and neck positions. Journal of Speech and Hearing Research. 2(3): 277 – 281.

Suter, A.H. 1989. The effects of noise on speech and warning signals. Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory. AD-A212 520.

Appendix A. Manufacturer's list.

Communications & Ear Protection, Inc.
PO Box 311174
3700 Salem Road
Enterprise, AL
36331-1174

Thales Acoustics
Waverley Industrial Estate
Hailsham Drive
Harrow, Middlesex
HA1 4TR, United Kingdom

Tucker-Davis Technologies
11930 Research Circle
Alachua, FL 32615

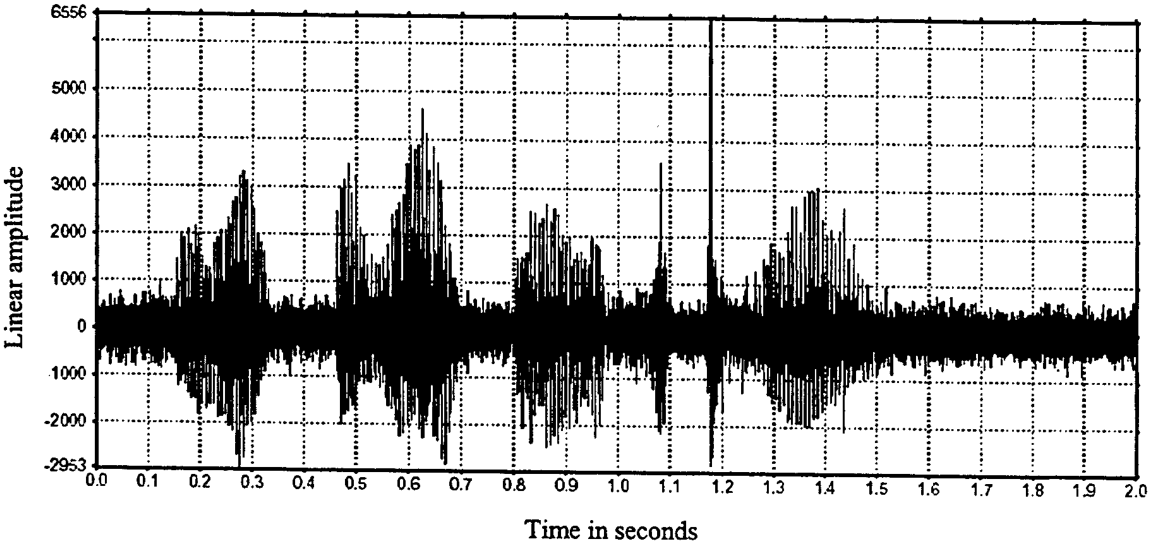
Appendix B. Sample waveforms and spectra.

Waveforms and spectra for the phrase "Select the word 'bent' please."

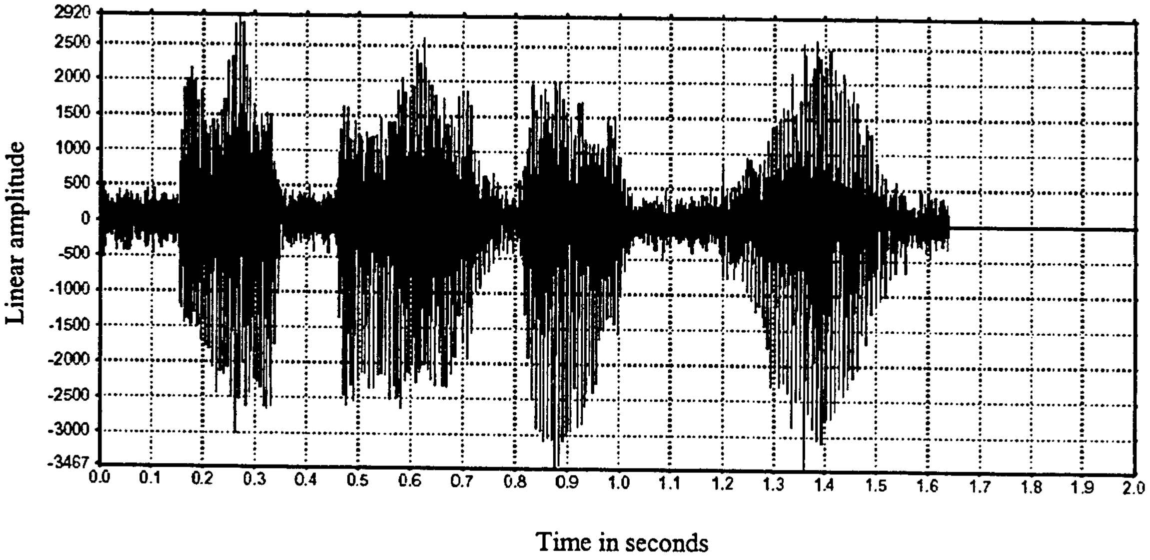
Time waveform

106 dB(A) recording noise

Acoustic boom microphone



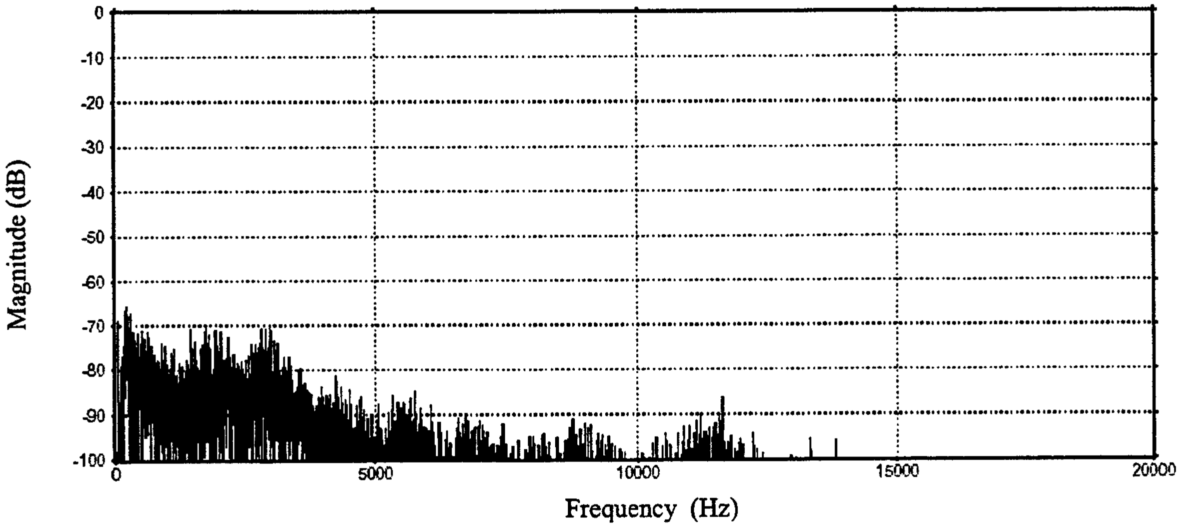
Throat microphone



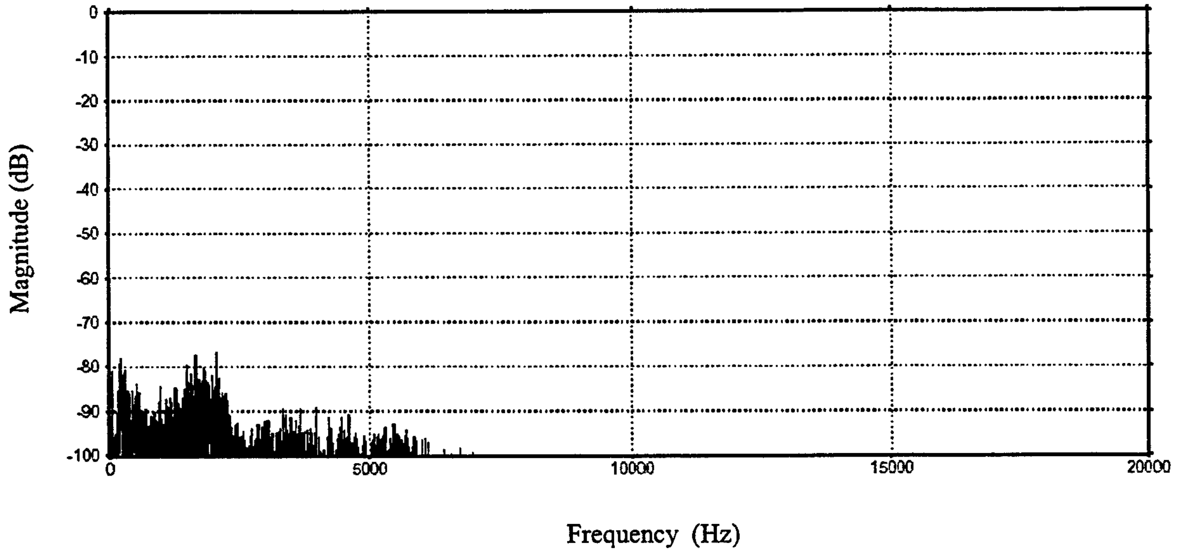
Frequency spectra

90 dB(A) recording noise

Acoustic boom microphone

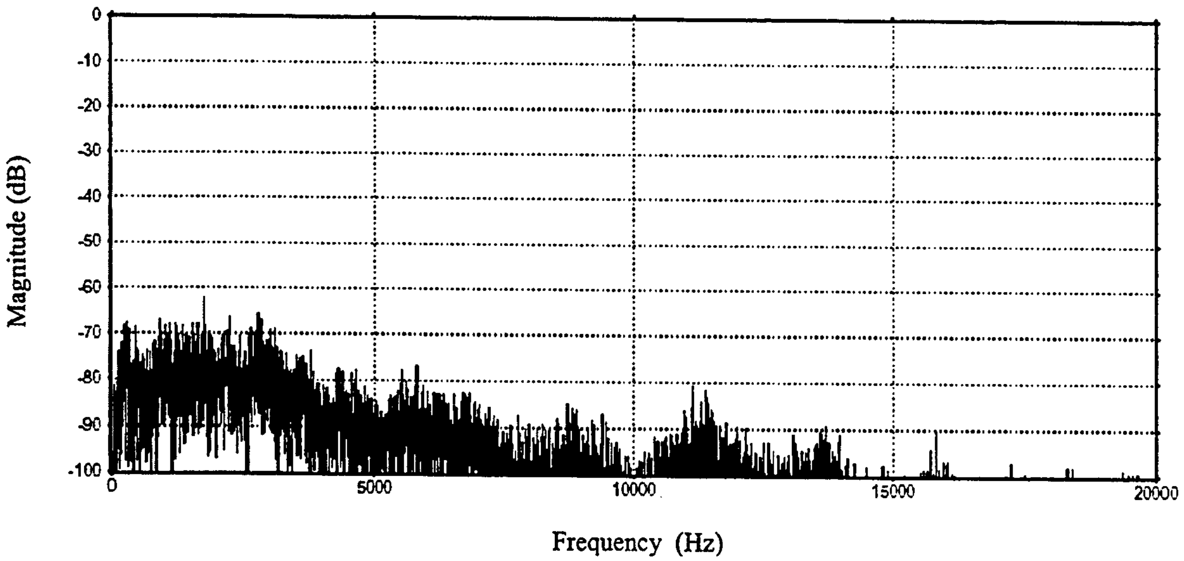


Throat microphone

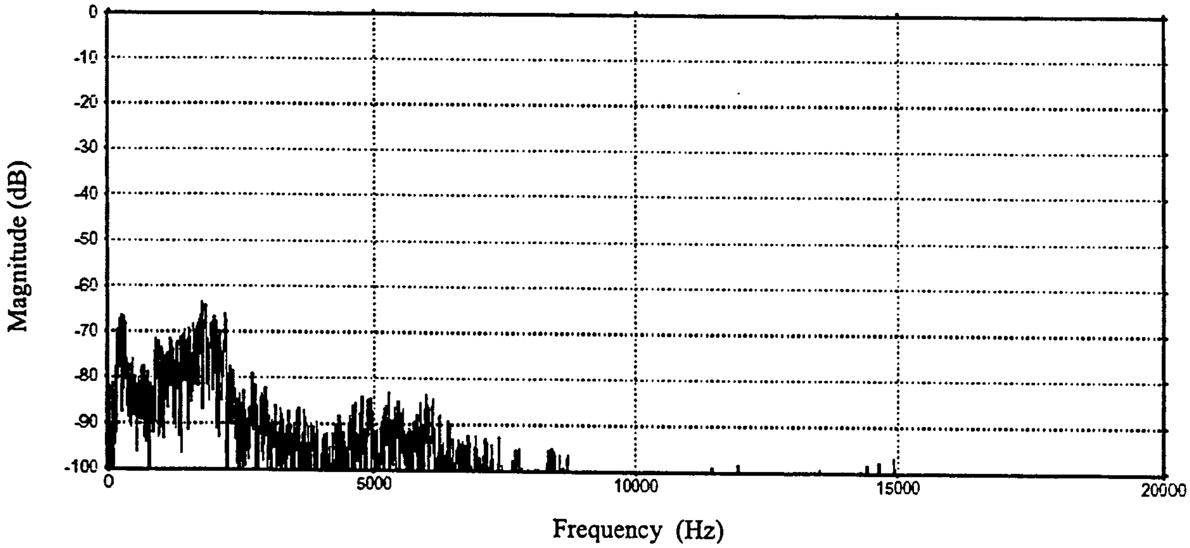


106 dB(A) recording noise

Acoustic boom microphone



Throat microphone



Appendix C. Analysis of variance summary table.

Effect	Sum of Squares	Degrees of Freedom	Mean Square	F	P
Microphone	4232.505	2	2116.253	32.215	.000
Recording noise	1948.424	1	1948.424	31.169	.000
Listening noise	5728.390	1	5728.390	16.303	.003
Microphone × Recording noise	580.120	2	290.060	5.232	.016
Microphone × Listening noise	.959	2	.479	.005	.995
Recording noise × Listening noise	17.222	1	17.222	.329	.580
Microphone × Recording noise × Listening noise	341.104	2	170.552	1.831	.189

Appendix D. Individual subject data.

Conditions are labeled according to the following three-digit code: Digit 1 = microphone type, Digit 2 = recording noise, and
Digit 3 = listening noise.

Microphone type: 1 = boom 2 = throat 3 = both

Recording noise: 1 = 90 dB(A) 2 = 106 dB(A)

Listening noise: 1 = 90 dB(A) 2 = 106 dB(A)

Subject	**111**	**112**	**121**	**122**	**211**	**212**	**221**	**222**	**311**	**312**	**321**	**322**
1	88.00	78.00	94.00	76.00	72.00	72.00	52.00	50.00	84.00	76.00	84.00	70.00
2	82.00	66.00	62.50	80.00	82.00	50.00	66.00	50.00	62.00	82.00	76.00	60.00
3	72.00	60.00	74.00	66.00	62.00	74.00	50.00	50.00	56.00	54.00	72.00	56.00
4	79.59	54.00	86.00	48.98	70.00	42.86	52.00	31.11	80.00	51.02	72.00	48.98
5	90.00	68.75	86.00	65.00	91.67	50.00	47.92	37.78	74.47	79.17	62.50	58.14
6	76.00	76.00	76.00	54.00	79.59	60.00	80.00	38.00	82.00	72.00	75.51	58.00
7	82.00	62.00	86.00	44.00	78.00	42.00	48.00	34.00	84.00	52.00	86.00	36.00
8	80.00	48.00	56.00	46.00	68.00	50.00	50.00	38.00	78.00	68.00	75.00	50.00
9	86.00	78.00	76.00	84.00	80.00	81.33	52.00	80.00	82.00	86.00	84.00	75.51
10	44.00	60.00	57.14	40.00	34.04	20.41	43.75	28.21	64.00	38.00	42.00	33.33