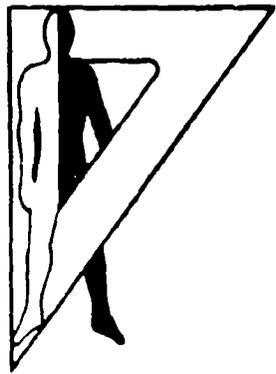


12

AD



Technical Memorandum 11-86

THE EFFECT OF HELICOPTER VIBRATION ON THE
ACCURACY OF A VOICE RECOGNITION SYSTEM

Thomas W. Dennison
Sperry Aerospace and Marine Group

Frank J. Malkin
Christopher C. Smyth
Human Engineering Laboratory

September 1986
AMCMS Code 612716.H700011

DTIC
ELECTE
NOV 17 1986
B

Approved for public release;
distribution is unlimited.

U. S. ARMY HUMAN ENGINEERING LABORATORY
Aberdeen Proving Ground, Maryland

AD-A174 284

DTIC FILE COPY

Destroy this report when no longer needed.
Do not return it to the originator.

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Use of trade names in this report does not constitute an official endorsement or approval of the use of such commercial products.

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|--|-------------------------------------|--|
| 1. REPORT NUMBER Technical Memorandum 11-86 | 2. GOVT ACCESSION NO. AD-A174284 | RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) THE EFFECT OF HELICOPTER VIBRATION ON THE ACCURACY OF A VOICE RECOGNITION SYSTEM | | 5. TYPE OF REPORT & PERIOD COVERED Final |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s) Thomas W. Dennison Frank J. Malkin Christopher C. Smyth | | 8. CONTRACT OR GRANT NUMBER(s) |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS Human Engineering Laboratory Aberdeen Proving Ground, MD 21005-5001 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBER AMCMS Code 612716.H700011 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS | | 12. REPORT DATE September 1986 |
| | | 13. NUMBER OF PAGES 13 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | | 15. SECURITY CLASS (of this report) Unclassified |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |
| 16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution is unlimited. | | |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) | | |
| 18. SUPPLEMENTARY NOTES | | |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Automatic Speech Recognition Helicopter Human Factors Aviation | | |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Speech recognition technology could be especially advantageous to single-crewmember helicopters like the Light Helicopter Family (LHX). Before speech recognition can be considered a viable technology for helicopters, several issues remain to be resolved. One of these is changes in the voice that occur as a result of stress, noise, and vibration. This paper reports the results of an investigation conducted to determine the effect of vibration-induced changes in the voice on the accuracy of a | | |

speech recognition system. A series of flight tests were conducted using 12 participants and 8 different flight maneuvers. Data were collected with the participant speaking 50 phonetically balanced words into the speech recognizer while seated in the copilot's seat of a UH-1H helicopter during each of the 8 flight maneuvers. The results indicate that speech recognition system accuracy is not affected by helicopter vibration.

S DTIC
ELECTE **D**
NOV 17 1986

B

QUALITY
INSPECTED
4

| | |
|--------------------|-------------------------------------|
| Accession For | |
| NTIS GRA&I | <input checked="" type="checkbox"/> |
| DTIC TAB | <input type="checkbox"/> |
| Unannounced | <input type="checkbox"/> |
| Justification | |
| By _____ | |
| Distribution _____ | |
| Availability Codes | |
| Avail and/or | |
| Dist | _____ |
| A-1 | |

THE EFFECT OF HELICOPTER VIBRATION ON THE
ACCURACY OF A VOICE RECOGNITION SYSTEM

Thomas W. Dennison
Sperry Aerospace and Marine Group

Frank J. Malkin
Christopher C. Smyth
Human Engineering Laboratory

Technical Assistance

Robert C. Brucksch
Human Engineering Laboratory

Paul Knight
Sperry Aerospace and Marine Group

September 1986

APPROVED:



JOHN D. WEISZ
Director
Human Engineering Laboratory

Approved for public release;
distribution is unlimited.

HUMAN ENGINEERING LABORATORY
Aberdeen Proving Ground, Maryland 21005-5001

CONTENTS

| | |
|---|----|
| INTRODUCTION | 3 |
| OBJECTIVE | 3 |
| METHOD | 4 |
| RESULTS | 7 |
| DISCUSSION | 8 |
| REFERENCES | 9 |
| APPENDIX | 11 |
| TABLES | |
| 1. Vibration Levels from Laing, Hepler, and Merrill (1973) | 6 |
| 2. Mean Transformed Recognition Accuracy Data by Group and Condition | 8 |

THE EFFECT OF HELICOPTER VIBRATION ON THE ACCURACY OF A VOICE RECOGNITION SYSTEM

INTRODUCTION

Computer recognition of speech is a potentially advantageous alternative to manual data input in helicopter cockpits. The use of voice allows pilots to keep their hands on the controls and their eyes focused outside the cockpit. This reduces manual and visual workload for input tasks performed concurrently with flying tasks (Dennison & Moore, 1985; Malkin & Christ, 1985a).

One of the issues that must be resolved before voice recognition can be considered a viable technology for helicopters is the effect of vibration-induced changes in the voice on the accuracy of a voice recognition system. Anecdotal reports suggest that the voice of helicopter pilots can sound "shaky" during some maneuvers. The French Service Technique des Telecommunications et Equipements Aeronautiques (STTE) tested a voice recognition system aboard French Navy and Air Force Puma helicopters and reported high (95% +) recognition rates ("France Completes," 1986). No details pertaining to the conditions of the study, like vibrations, were reported. Dennison (unpublished, 1985) used an 8000-pound capacity "vibration table" with a Bell Cobra seat mounted to it to expose subjects to various levels of vertical vibrations. Subjects enrolled a 50-word vocabulary while not vibrating; then spoke the words at baseline (no vibration) and all six combinations of vibration frequencies of 6 Hz and 24 Hz and accelerations of .05g, .075g, and .1g. The frequencies represent the one-per-rev and the four-per-rev harmonics of a four-bladed rotor system. No difference in recognition accuracy was detected between baseline and any of the vibration conditions. Cruise, Denson, and Rajasekaran (in press) obtained similar results using a vibration table with a different voice recognition system.

Studies performed with a vibration table allow precise control of vibration level, but they lack realism. The current study makes use of an Army UH-1H helicopter to evaluate the effects of realistic, in-flight vibration levels. Other differences between the studies are that (1) the UH-1H has a two-bladed rotor system while the vibration levels selected in the "shaker table" studies simulated a four-bladed rotor system and (2) subharmonics that are not produced by the vibration table are present in flight.

OBJECTIVE

The objective of this study was to conduct a flight test to examine the effect of vibration-induced changes in the voice on the accuracy of a voice recognition system.

METHOD

Participants

Twelve males volunteered to serve as subjects. Six of the participants were civilian employees of the Human Engineering Laboratory, Aberdeen Proving Ground, Maryland. The remaining six were U.S. Army aviators.

Apparatus

All data were collected with participants seated in the copilot's seat of a U.S. Army UH-1H helicopter. A B&K ride pad accelerometer and amplifiers for each axis were employed to record vibration data. Participants used a noise-canceling microphone attached to a binaural headset. A Votan VPC-2000 was integrated with an IBM PC XT. This was repackaged into a flightworthy unit by SCI Systems. The Votan VPC 2000 is a speaker-dependent voice recognition system which requires a sample of how each user pronounces the utterance in a predetermined vocabulary prior to use. This is referred to as training or enrolling the system. Each sample is stored in memory as a reference for later comparisons. When in use, the system recognizes words by comparing current utterances with the samples stored in memory and selecting the closest match. The acceptance setting for the recognizer was left at the default value of 50 on a scale of 1-255; the gain was set at 2 on a scale of 1-7.

Participants' utterances, vibration data, cockpit noise levels, and aircraft communications were recorded using a Honeywell 6300E 28-channel data recorder. Cockpit noise levels were measured with a B&K noise level meter. The vocabulary consisted of 50 words (see Appendix) from the Phonetically Balanced Word List Speech Intelligibility Test, U.S. Army TOP 1-2-610. The occurrence of phonemes in this list is proportionate to their occurrence in English. A tape loop of recorded UH-1H noise was played on a Nagra tape recorder through loudspeakers placed in an acoustic reverberation chamber. The baseline and all in-flight recognition data were recorded on 5.25-inch, double-sided, double-density floppy disks.

Procedure

Noise can have a detrimental effect on speech recognition system performance and it is a potential confounding effect. Previous research has shown that if a speech recognition device is trained in a quiet environment followed by attempts to use it in a noisy environment, severe degradation in performance can result. When the device is trained and used in a noisy environment, recognition accuracy is the same as if the device had been trained and used in a quiet environment (Kersteen, 1982). In order to control the effects of noise, the enrollment of the voice recognition system was conducted in the presence of helicopter noise.

Training

Participants were trained and tested individually. Each participant was briefed on the purpose of the experiment and given general instructions on its conduct. Training took place in the acoustic reverberation chamber at the acoustics laboratory at the Human Engineering Laboratory. The participant was familiarized with the Votan system and practiced using the recognizer with a sample six-word vocabulary until it was evident that the test procedure and use of the voice recognition system were clearly understood. Next, the 50-word vocabulary was enrolled into the voice recognition system. The words were read from a list, and the list was repeated three times as recommended by the manufacturer. The enrollment was accomplished in the presence of taped helicopter noise. Participants wore ear protection and headsets during the enrollment.

Testing

The baseline (no vibration) data were also collected in the acoustic reverberation chamber in the presence of taped helicopter noise. The participant read the 50-word list. A short beep was presented through the participant's earphones as a cue to say the appropriate word. The words were spoken in order, and the participant was instructed to wait for the beep before speaking the next word. From there, the participant and the voice recognition box were moved to the airfield. Data trials were accomplished with the participant seated in the copilot's seat. The maneuver conditions tested in the helicopter were:

- 500-feet per minute climb
- hover-out-of-ground effect
- 45-degree bank turn
- 60-knot level flight
- 110-knot level flight
- 500 feet per minute descent
- hover-in-ground effect
- ground idle

When the pilot signaled that he had established the appropriate maneuver, the participant was signaled to begin reading the list. As in the baseline data collection, the participant was instructed to wait for the beep prompt before saying each word. Although it was impractical to counterbalance the maneuver conditions, alternate presentation orders were used to test for ordering effects.

The vibration levels that occur in a UH-1H helicopter under the maneuver conditions are shown in Table 1. These data indicate the frequency in Hertz and the intensity in g's of the first three harmonics (Laing, Hepler, & Merrill, 1973). The data were gathered using an accelerometer mounted on the pilot's seat to record whole body vibrations. An attempt was made to record vibration data during the maneuver conditions of the present study; however, the data were lost.

TABLE 1

Vibration Levels from Laing, Hepler, and Merrill (1973)

| | Harmonics | | |
|--------------|---------------|---------------|---------------|
| | 1st Hz/g's | 2nd Hz/g's | 3rd Hz/g's |
| Level Flight | 7.5/.08 | 20/.05 | 27/.04 |
| Hover | 10/.03 | 17.5/.03 | 27/.02 |
| Climb | 6/.03 | 10/.08 | 22.5/.04 |
| Descent | 5/.03 | 10/.07 | 20/.04 |
| Bank | 5/.06 | 10/.1 | 20/.01 |
| Grid Idle | 5/.03 | 17.5/.02 | 25/.01 |

RESULTS

Because two subject populations were used (aviators and nonaviators), the two groups were factored into the experimental design. It was expected that nonaviators may have encountered a higher degree of stress during the flight maneuvers than aviators. Stress has been shown to affect vocal output (Dennison & Moore, 1984) and could have confounded the results.

Two participants' data were dropped because a large number of their utterances were not processed by the recognizer. This was an intermittent problem that occurred with one aviator and one nonaviator, and it did not occur in all conditions. The software written to automate the data collection required a slight (750-millisecond) pause after the beep prompt during which the recognizer would sample the utterance. If the participant spoke the word too quickly, it would not be sampled and recognized by the processor.

Raw data, in percent, for each condition were converted using the arcsine conversion (Ferguson, 1976) to satisfy the assumptions of the Analysis of Variance.

$$X_1 = \arcsine \sqrt{X}$$

where:

X = raw % score

X₁ = converted score

The raw data are not reported because the intent of this study is to compare performance between baseline and maneuver conditions; the intent is not to compare the performance of this recognition system with that of other systems. Such a comparison would constitute a misapplication of data unless the performance of the two systems had been measured under the same conditions. Table 2 shows the means for each condition and submeans for the five aviators versus five nonaviators.

An analysis of variance with repeated measures for the maneuver conditions was conducted. The nonrepeated factor was aviators versus nonaviators. No significant difference was detected between the recognition accuracy of aviators versus nonaviators over all levels of maneuver conditions. Likewise, no significant differences were detected among the mean accuracy scores of any of the maneuver conditions, including the baseline condition. No significant interaction was noted; however, a significant between-subjects effect was found, $F(8,64)=3.63$, $p<.01$.

TABLE 2

Mean Transformed Recognition Accuracy Data in Percent
by Group and Condition

| Group | <u>Conditions</u> | | | | | | | | |
|------------------|-------------------|------|------|------|------|------|------|------|------|
| | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 |
| Aviator Means | 60.4 | 56.6 | 53.5 | 57.3 | 63.0 | 54.1 | 53.8 | 57.0 | 57.0 |
| Nonaviator Means | 64.3 | 54.7 | 51.9 | 48.3 | 53.3 | 55.0 | 49.7 | 55.9 | 56.7 |
| Grand Means | 62.4 | 55.6 | 52.7 | 52.8 | 58.2 | 54.5 | 51.7 | 56.4 | 56.8 |

C1 - baseline (no vibration)

C2 - 500-feet per minute climb

C3 - hover-out-of-ground effect

C4 - 45-degree bank turn

C5 - 60-knot level flight

C6 - 110-knot level flight

C7 - 500 feet per minute descent

C8 - hover-in-ground effect

C9 - ground idle

DISCUSSION

Whether a participant was an aviator or a nonaviator had no effect on the accuracy of recognition of his utterances. The possibility that this variable confounded the results can, therefore, be ruled out. More importantly, there was no difference in recognition accuracy between the baseline condition and any of the in-flight maneuver conditions, nor were there any differences among the maneuver conditions. These findings agree with those of the previously cited studies (Dennison, unpublished, 1985; Cruise, Denson, & Rajasekaran, in press) conducted on "shaker tables" in that vibration caused no degradation of recognition system accuracy. Moreover, these results go beyond those obtained in simulated vibration conditions in providing real-world verification that vibration will not impede the use of voice recognition technology in helicopters. Of interest would be a study that explores the limits of vibration at which recognition will be degraded. It may be difficult, however, to perform this research without endangering subjects. The possibility of computer modeling of the vocal system and its response to vibration should be explored.

The significant between-subjects effect indicates that there were greater than chance differences among the participants' recognition rates when compared to each other. Other studies have reported high variability in recognition rates based on individual differences (Aretz, 1983; Malkin, 1983; Malkin & Christ, 1985b)

This may be one of the most urgent problems that users and makers of speech recognition algorithms face. If speech recognition systems are to be acceptable for extensive use in Army helicopter cockpits, then manufacturers must provide systems that recognize a wider range of the population.

REFERENCES

- Aretz, A.J. (1983). A comparison of manual and vocal response modes for the control of aircraft subsystems (AFWAL-TR-83-3005). Wright-Patterson Air Force Base, OH: Flight Dynamics Laboratory.
- France completes tests of voice unit on Puma. (1986, January 6). Aviation Week & Space Technology, 124(1), 27.
- Cruise, R., Denson, D., & Rajasekaran, P. (in press). Voice-recognition in the helicopter environment.
- Dennison, T., & Moore, L. (1985). The effect of voice versus manual input on input speed and external visual tracking accuracy (Task I Report). Fort Worth, TX: Bell Helicopter Advanced Rotorcraft Technology Integration Program.
- Dennison, T., & Moore, L. (1984, October). The effect of operator workload on the robustness of speech recognition algorithms. Paper presented at American Helicopter Society Specialists' Meeting, Dallas, TX.
- Dennison, T. (1985). The effect of simulated helicopter-vibration on the accuracy of a voice recognition system. Unpublished manuscript.
- Ferguson, G. (1976). Statistical analysis in psychology and education. New York: McGraw Hill Company.
- Kersteen, Z. (1982). An evaluation of automatic speech recognition under three ambient levels. Proceedings on Workshops on Standardization for Speech I/O Technology. Gaithersburg, MD: National Bureau of Standards.
- Laing, E.J., Hepler, L.J., & Merrill, R.K. (1973). Vibration and temperature survey: Production UH-1H helicopter (Project No. 70-15-2). St. Louis, MO: U.S. Army Aviation Systems Command.
- Malkin, F.J. (1983). The effects on computer recognition of speech when speaking through protective masks (TM 7-83). Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory.
- Malkin, F.J., & Christ, K.A. (1985a). A comparison of voice and keyboard data entry for a helicopter navigation task (TM 17-85). Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory.
- Malkin, F.J., & Christ, K.A. (1985b). Human factors assessment of voice technology for the light helicopter family (LHX) (TM 9-85). Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory.
- U.S. Army Test and Evaluation Command. (1983, November). U.S. Army TOP 1-2-610. Aberdeen Proving Ground, MD: Author.

APPENDIX
VOCABULARY

VOCABULARY

| | |
|--------|--------|
| FRIGHT | LOUSE |
| TURN | PITCH |
| AID | PUMP |
| WIELD | CREWS |
| GAB | TUCK |
| ROUGE | TON |
| DUMP | ROCK |
| MAP | SUIT |
| HOSE | DAME |
| STRESS | TIRE |
| RUG | VOW |
| BOOK | SHEEP |
| LEASH | STAB |
| CLIFF | INK |
| FIFTH | SORE |
| THRESH | THREE |
| BARGE | DUB |
| LAY | RYE |
| DIN | CHEESE |
| SHIEK | KIND |
| PART | NEXT |
| HAD | CLOSED |
| SANG | GAS |
| KNEE | DRAPE |
| HASH | NAP |