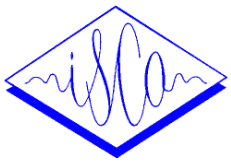

ESCA-NATO Workshop on
Speech under Stress
Lisbon, Portugal
September 14-15, 1995

Voice Changes During Sustained Duty Periods

Jeffrey Whitmore and William Storm, Ph.D.

Armstrong Laboratory/Sustained Operations Branch

ABSTRACT

Research was conducted to determine if alterations in the acoustical characteristics of voice occur over periods of sustained wakefulness. Twelve male United States Air Force B-1B aircrewmen participated in the study. The participants served in teams of four and performed three 36-hour experimental periods (missions) in a high-fidelity simulator. The missions were interspersed with 36-hour rest breaks. Data were lost from two members of the third team due to a communication malfunction. Speech, cognitive, and subjective fatigue data were collected approximately every three hours for 11 trials per mission. Fundamental frequency and speech duration were both found to vary significantly over trials (fundamental frequency $F(10,90) = 2.63$, $p=.0076$, speech duration $F(10,90) = 2.5$, $p=.0106$). Speech duration results also showed a significant main effect of mission ($F(2,18) = 6.91$, $p=.0082$). The speech data follow the same trend as the data from the cognitive tests and subjective measures. A strong diurnal pattern is reflected in nearly all of the dependent measures. Overall, the results support the proposition that voice may be a reliable and valid indicator of a speaker's fatigue state.

INTRODUCTION

Voice has been shown to be a reliable indicator of internal state. Williams and Stevens (1969) examined the emotional state of pilots during flight and found significant vocal fluctuations under conditions of stress. Pisoni (1990) has shown some extreme environments (increased gravity, noise, increased cognitive workload) to influence vocal characteristics. Increased stress or workload has also been shown to affect voice (Griffin & Williams, 1987; Brenner & Shipp, 1988). As well, intoxication effects upon voice are well documented (Brenner & Cash, 1991; Klingholz, Penning & Liebhardt, 1988). In general, increases in fundamental frequency and amplitude, and decreases in speech duration have been observed in voice when the speaker is involved in a stressful activity. The obverse is true for intoxication, with fundamental frequency decreasing and speech duration increasing.

The purpose of the following work was to examine voice recordings taken over an extended time to determine the effect of fatigue on simple voice characteristics. Fundamental frequency and speech duration measures were taken over three 36-hour simulated missions. Subjective fatigue and cognitive measures were also taken to provide a baseline for the voice measures to be compared against. Overall, fundamental frequency was expected to decrease and speech duration to increase as fatigue increased in the speaker.

METHOD

This study was a simple repeated measures design. Each participant participated with his flight crew in a single 10 day session. This session consisted of one day of practice on the cognitive tests followed by three repetitions of a 36 hour on, 36 hour off cycle and concluded in one final day for recovery. Participants arrived at crew quarters at 10 a.m. the day prior to the first simulated mission. The following day they arrived at the weapons system training facility at 8 p.m. and remained there for approximately 36 hours. They were then allowed 36 hours of rest returning again to the simulator at 8 p.m. the following day (Fig 1).

Prep	Flight	Rest	Flight	Rest	Flight	Rest
24 Hr	36 Hr	36 Hr	36 Hr	36 Hr	36 Hr	24 Hr

Figure 1

Voice samples, subjective fatigue and cognitive performance scores were taken every three hours throughout each of the simulated flights. All voice samples were recorded under the same conditions including approximately the same amount of system

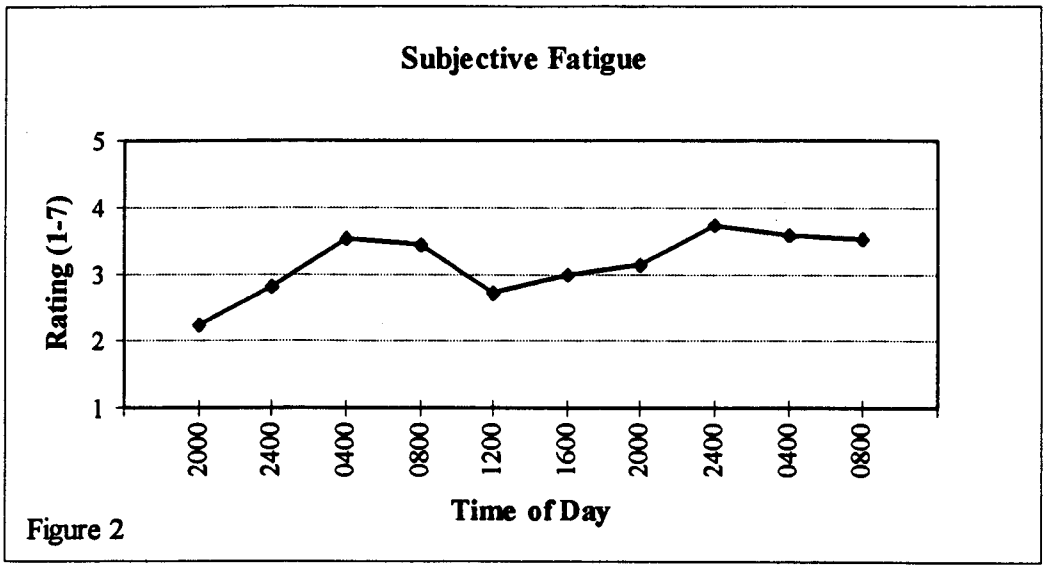

Figure 2

noise. Voice was measured through the use of the phrase "Futility Magellan. This is 'name'. The time is 'time'." Both voice measures were calculated on the word "Magellan." Subjective fatigue was taken on a 7-point Likert scale with 1 being low fatigue and 7 being extreme fatigue. Cognitive performance was measured using a computerized test battery consisting of the following tests: Logical Reasoning, Tracking, Matrix Comparison, and Attention Switching. Accuracy (Acc) measured as percent correct and reaction time (RT) to correct responses were measured for each test.

RESULTS

An alpha level of .05 was used for all statistical tests. A Two-Way Repeated Measures ANOVA was performed on the fatigue data. There was no significant interaction between missions and trials. The effect of missions was not significant. Subjective fatigue decreased for each successive mission, from 3.4 for the first to 3.2 for the second and 3.0 for the third. There was a significant main effect of trials ($F(9,81) = 5.54$, $MSE = 1.11$). Fatigue data for each mission showed a strong circadian cycle (Fig 2)

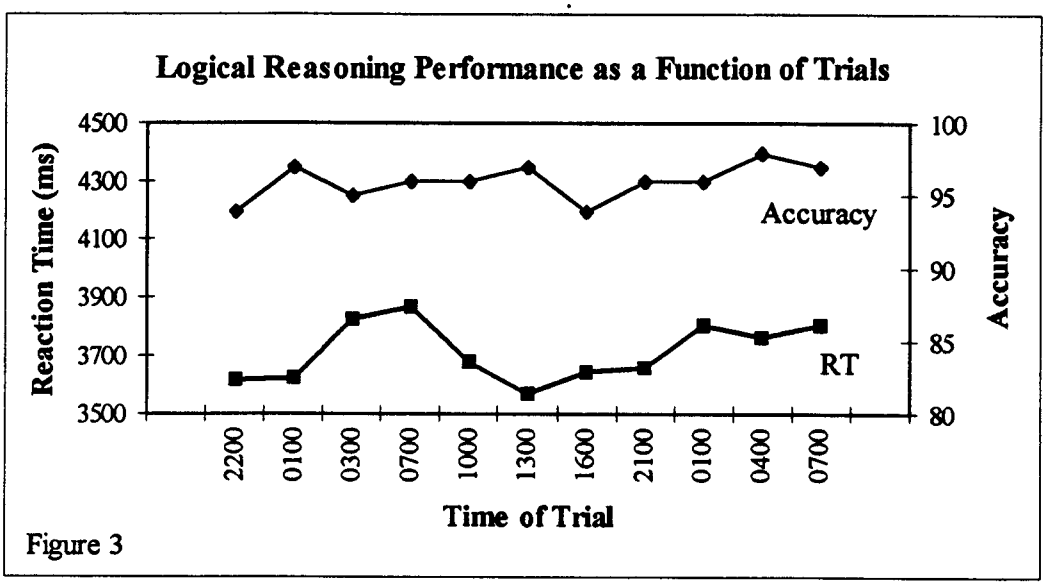

Figure 3

The Logical Reasoning Task results provide an example of the performance on the computer tests (Fig 3). In general accuracy varied only slightly over trials, while RT showed greater and more consistent

variation. A circadian pattern was seen in each of the RT measures. A Two-Way Repeated Measures ANOVA was performed for accuracy and RT on each cognitive test. The significance tests results are shown in Figure 4. There was no interaction of missions and trials for any of the measures.

A Two-Way Repeated Measures ANOVA was performed on the fundamental frequency data. There was no significant interaction between trials and missions and no significant effect of missions. There was a significant main effect of trials ($F(10,90) = 2.63$, $MSE = 70.59$). The results of the analyses on fundamental frequency were consistent with a circadian pattern (Fig 5).

The interaction between trials and missions was not significant for speech duration. There was a significant main effect of missions ($F(2,18) = 6.91$, $MSE = .0054$). There was also a significant main effect of trials ($F(10,90) = 2.50$, $MSE = .0026$). Speech rate averages were within the expected range of 4.5 to 7.5 syllables per second. The circadian trend present in all the earlier RT results was again present in the speech duration data (Fig 6).

COGNITIVE TESTS RESULTS

TASK	MEASURE	MAIN EFFECTS	
		T- significant trial effect M - significant mission effect	
Logical Reasoning			
	RT	T $F(10,90) = 3.62$	$p=.0030$
	Acc	T $F(10,90) = 2.37$	$p=.0130$
		M $F(2,18) = 7.28$	$p=.0031$
Unstable Tracking			
	Hits	T $F(10,90) = 4.23$	$p=.0001$
		M $F(2,18) = 8.46$	$p=.0026$
Continuous Recognition			
	RT	T $F(10,90) = 6.25$	$p=.0001$
		M $F(2,18) = 15.31$	$p=.0003$
	Acc	no significant effects	
Attention Switching Task			
	Mannequin		
	RT	T $F(10,90) = 4.12$	$p=.0001$
		M $F(2,18) = 4.28$	$p=.0301$
	Acc	T $F(10,90) = 1.96$	$p=.0460$
	Math		
	RT	T $F(10,90) = 5.77$	$p=.0001$
		M $F(2,18) = 5.8$	$p=.0114$
	Acc	no significant effects	

Figure 4

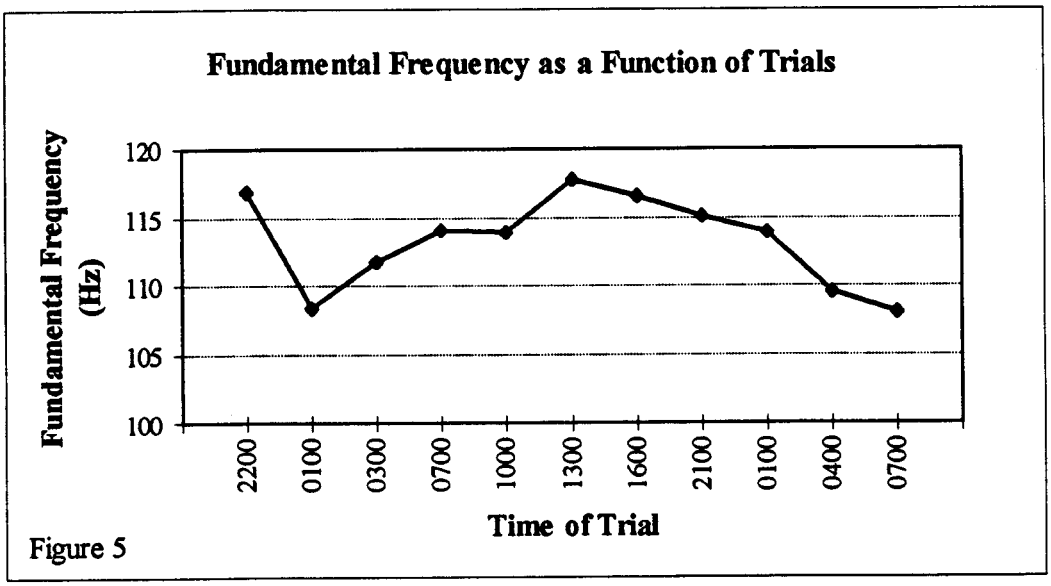

Figure 5

CONCLUSION

One of the most consistent findings in this study is that cognitive performance tended to increase over missions, not decrease. Practice and adaptation to the somewhat novel environment of the simulator were the factors most likely responsible for changes in performance over the three missions.

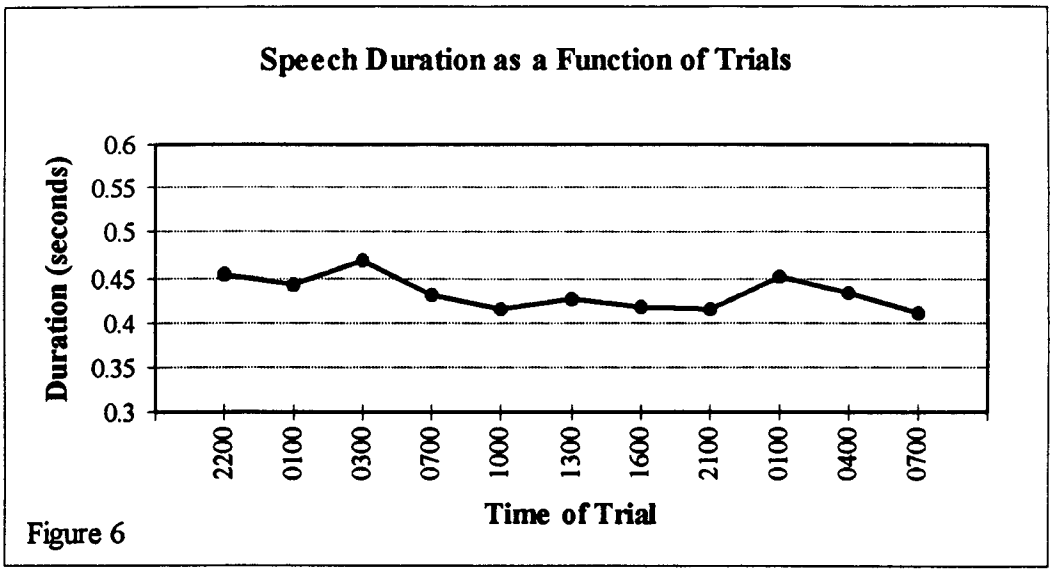

Figure 6

Within each mission the normal circadian cycle was present, where cognitive performance was lowest during the early A.M. hours. The voice measures paralleled these changes. A quartic trend analysis was performed on each of the RT measures and speech measures. This test was significant for all reaction time measures except Logical Reasoning. Both fundamental frequency and speech duration returned significant values on this test as well.

It appears that some vocal properties do indeed fluctuate like subjective fatigue and cognitive performance over periods of sustained wakefulness. The circadian cycle seen in the cognitive test RT data is also present in the speech duration data and the fundamental frequency data. The contents of post-mission interviews with the aircrews contained no complaints of hoarseness or vocal strain. Thus, it is reasonable to assume that any changes to the voice signal resulted from the effects of sustained wakefulness, not any acute situation of speaking loudly or speaking non-stop for a lengthy period.

REFERENCES

Brenner, M., & Cash, J. R. (1991). Speech analysis as an index of alcohol intoxication-The Exxon Valdez Accident. Aviation, Space, and Environmental Medicine, 62, 893-898.

Brenner, M., & Shipp, T. (1988). Voice Stress Analysis: Mental-State Estimation. In J. R. Comstock, Mental-State Estimation 1987 (pp. 363-376). (NASA 2504).

Griffin, G. R., & Williams, C. E. (1987). The effects of different levels of task complexity on three vocal measures. Aviation, Space and Environmental Medicine, 58, 1165-1170.

Klingholz, F., Penning, R., & Liebhardt, E. (1988). Recognition of low-level alcohol intoxication from speech signal. Journal of the Acoustical Society of America, 84(3), 929-935.

Pisoni, D. B. (1990). Speech Perception and Production in Severe Environments. AAMRL-SR-90-507.

Williams, C. E., & Stevens, K. N. (1969). On determining the emotional state of pilots during flight: an exploratory study. Aviation Medicine, 40, 1369-1372