Acoustic analysis of the effects of 24 hours of sustained wakefulness

Adam P. Vogel [1], *Janet Fletcher* [2], *Paul Maruff* [1]

[1]Centre for Neuroscience, University of Melbourne, Australia
[2] Linguistics and Applied Linguistics, University of Melbourne, Australia
`avogel@cogstate.com, janetf@unimelb.edu.au, pmaruff@cogstate.com`

Abstract

The effect of 24 hours of sustained wakefulness on the speech of healthy adults is poorly documented. Therefore, speech samples were systematically acquired (e.g., every four hours) from 18 healthy adults over 24 hours. Stimuli included automated and extemporaneous tasks, sustained vowel and a read passage. Measures of timing and frequency were derived acoustically using Praat and significant changes were observed on all tasks. The effect of fatigue on speech was found to be strongest just before dawn (after 22 hours). Key features of timing (e.g., mean pause length), frequency (e.g., F4 variation) and power (alpha ratio) changed as a function of increasing levels of fatigue.

Index Terms: fatigue, voice, tiredness, clinical marker

1. Introduction

The management of fatigue is an important issue in society and there is increasing interest in the development of objective non-invasive systems that can be used to assist the identification and management of fatigue in both health and workplace settings (e.g., military operations, extended hours performed by medical workers) [1]. Assessment of the acoustic properties of speech provides reliable insight into damage and injury to the central nervous system (CNS) and even in healthy adults, changes in acoustic properties occur in response to environmental, physical and pharmacological challenges. It is well accepted that speech changes as healthy individuals become physically fatigued with studies showing changes in timing [2] and pitch [3].

Although changes in speech have been observed in fatigued individuals, there does not exist sufficient data regarding the nature or magnitude of changes in acoustic properties because previous studies have been based on small sample sizes [4], use poor quality recording technology [5], utilize limited or subjective outcome measures [2] and rely on periods of sustained wakefulness greater than 34 hours, suggesting that changes in speech were not anticipated in periods shorter than this. Furthermore because the identification of fatigue related changes in acoustic properties requires that the same individuals undergo assessment when they are both not fatigued and fatigued, it is important that the speech tasks chosen and the outcome measures generated are statistically reliable, are stable in the absence of any true CNS change and for which the repeated application itself does not engender change (i.e. as with practice effects) [6]. Left uncontrolled these statistical and practical issues can act to obscure true change and thereby diminish the sensitivity of putative acoustic markers of fatigue.

The speech-fatigue literature describe several components of prosody that change as a function of increasing levels of fatigue, including variations in pitch, speaking rate and spectral energy. Specifically, reduced intonation ("monotonic or flattened voices" p871) [3] was observed perceptually during a counterbalanced experiment over 34 hours where 10 untrained listeners were asked to rate speech samples acquired from nine participants before and after 36 hours of sustained wakefulness. Research conducted with military personnel have documented changes in broad measures of timing including word duration [5] and total signal time [2]. Recently, Greeley et al. [4] looked at the first four formants (F1-F4) derived from a list of words that were produced at regular intervals over 36 hours. The authors used Sleep Onset Latency (SOL) [7] and reaction time to determine levels of fatigue and found these correlated with a limited number of formants (only F3 on /o/ and /u/ phones out of 76 candidates revealed strong correlations (≥0.7) with reaction time).

Evidence from the vocal fatigue (e.g., spectral tilt) and sustained wakefulness (e.g., timing and frequency) literature shows that speech changes can occur in different experimental models of fatigue, yet, the magnitude and nature of these changes has not been systematically documented using acoustic methodologies. Nor have objective changes in speech been considered using a within person design coupled with appropriate statistical procedures designed to increase sensitivity [6]. Therefore, the current study aimed to apply a methodology that objectively monitors the fatigue of individuals by repeatedly capturing changes in acoustic output over time, thus providing insight into overall CNS integrity.

2. Methods

2.1 Data Acquisition

The speech battery was performed by each participant individually in a quiet room. Speech samples were recorded using a laptop PC (Hewlett-Packard, Palo-Alto, CA) with basic factory settings and a Logitech A-0374A USB Headset uni-directional head-mounted microphone (Logitech, China) (minimum sensitivity of -38 dB and a frequency range of 100 Hz – 1 kHz) which was positioned at a 45 degree angle 8 cm from the mouth. All data was sampled at 44.1 kHz, coupled with quantization at 16 bits. Data was recorded using Audacity.

2.2 Participants

Eighteen (11 male, 7 female) healthy adult participants aged between 18 and 28 years ($\bar{x}$ age 22.7 ± 2.9) were recruited for the study. Participants were excluded from the study if they were heavy smokers, heavy coffee drinkers, recreational drug or alcohol abusers, participants with neurological trauma (determined via self report), or presenting with poor vocal health.

ISBN 978-0-9581946-3-1 © 2010 ASSTA
Accepted after peer review of full paper

2.3 Stimuli

Participants were instructed to speak in a natural manner, using their typical speaking rate, with a vocal effort appropriate for speaking to one or two people in quiet surroundings. The speech battery consisted of: 1) reading a phonetically balanced text, the grandfather passage (GRAN) (178 syllables); 2) sustain an open vowel /a:/ for approximately six seconds (AAAH); 3) counting from 1-20 (C120); 4) saying the days of the week beginning with Monday (DAYS) using one breath; and 5) extemporaneous speech task requiring participants to produce a monologue with a positive subject manner (i.e., happy memory, amusing story, topic of interest to participant) (FREE) for approximately one minute.

2.4 Procedure

The entire speech battery was elicited three times in succession during the first recording session at 08:00 in accordance with recommendations outlined in Vogel et al. [6], in order to acquire a stable production. The speech battery was then repeated once every four hours until midnight (at midday, 16:00, 20:00 and midnight). The battery was then repeated every two hours until 08:00 (at 02:00, 04:00, 06:00 and 08:00). The condensed recording regime was employed after midnight to capture any anticipated acoustic changes that may occur when the effect of sustained wakefulness was greatest. Participants remained on the premises from 07:30 until 08:00 the next day.

2.5 Acoustic Analysis

Each speech sample was segmented and analyzed using Praat (version 5.0.46) [8]. Silences were removed from the start and end of the AAAH task, C120, DAYS and GRAN passages. The AAAH and FREE samples were truncated 1.5 and 20 seconds respectively each side of the temporal midpoint. Segmentation produced samples with uniform signal lengths, allowing automated analysis of each sample.

2.5.1 Frequency

f0, F1-F4 and their corresponding standard deviations (SD) and coefficients of variances (CoV) were derived from the AAAH task using an automated Praat script [9]. *f0*, *f0* SD and *f0* CoV measures were taken from the FREE, C120, DAYS and GRAN tasks.

2.5.2 Timing

Timing measures were derived from the FREE, GRAN and automated speech tasks (C120, DAYS) using a modified version of the methodology outlined in Rosen et al. [10]. Three thresholds were defined to identify pauses from the intensity contour: (a) intensity threshold (0.65 of the distance between the peak intensity (peak intensity equaled 0.95 of the maximum intensity) and floor (i.e. minimum) intensity), (b) minimum pause duration (15ms), and (c) minimum speech duration (30ms).

2.5.3 Intensity

Power was calculated by comparing intensity at the higher end of the spectrum (above 1kHz) with intensity at the lower end of the spectrum (below 1kHz) (alpha ratio [11]) and by calculating the average sound level along the frequency axis, reflecting both glottal and vocal tract characteristics [12]. Given that vocal loudness is thought to affect LTAS slope, the overall equivalent sound level (L_{eq}) was calculated [13].

2.6 Statistical analysis

Data from each outcome measure for each subject were submitted to a series of ANOVAs (with repeated measures on assessment time: baseline, 8, 12, 16, 18, 20, 22, 24 hours). No significant differences were observed on ANOVAs between assessments on all acoustic measures. The magnitude of change from baseline was computed between the baseline (average of third recording at 08:00 and sole recording made at 12.00) recording and each post-baseline recording (16:00, 20:00, midnight, 02:00, 04:00, 06:00, 08:00) using a series of dependent groups t-tests. The t-statistic and correlation from the paired t-tests were then used to express the magnitude of the difference between baseline and each post-baseline time-point using the formula for repeated measures designs Dunlap's d [14].

3. Results

Figures 1 to 4 show the magnitude of change from baseline for frequency, timing and intensity expressed as effect size. Only those acoustic measures with noteworthy changes from baseline have been displayed and the time points reaching significance are indicated with an asterisk (* = p<0.05).

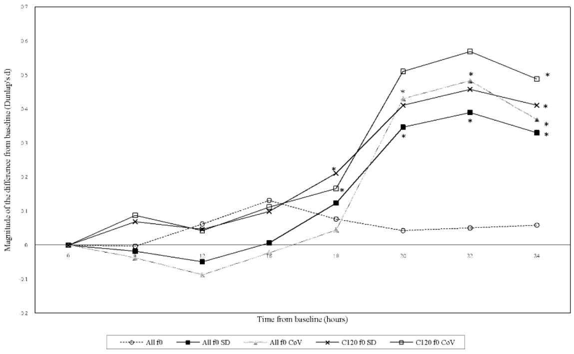

Figure 1: Magnitude of the difference (Dunlap's d) in change from baseline performance of frequency measures from the C120 task and all tasks excluding AAAH over 24 hours.

3.1 Frequency measures

Figure 1 shows the combined data for all tasks (except AAAH) on measures of *f0*, *f0* SD and *f0* CoV. All measures remained quantitatively stable for approximately 12 hours from baseline, at which point they began to increase. Between 16 – 24 hours post-baseline assessments, *f0* variation was significantly greater than baseline recordings. For measures of *f0* variation, the magnitude of the increase from baseline was reduced at the last recording session (sunrise), as the scores moved toward baseline levels. Figure 2 shows the effect of sustained wakefulness on F4. The greatest change from baseline was evident between 20 and 22 hours post-baseline. A clear return to baseline functioning is evident at 24 hours post-baseline on F4 SD/CoV.

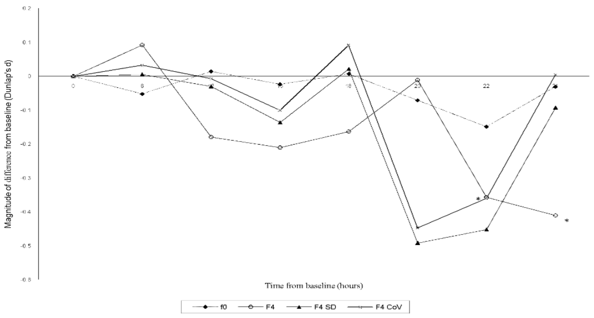

Figure 2: Magnitude of the difference (Dunlap’s d) in change from baseline performance of *f0* and formant variation of F4 on the AAAH task over 24 hours.

3.2 Timing measures

Figure 3 shows significant increases in total speech time, mean pause length (ms), total signal time and pause length SD as well as a corresponding reduction in speaking rate as time post baseline increased on the GRAN task. Increases in the speech to pause ratio at 20 and 22 hours post baseline, as well as a return towards baseline productions were also observed on the DAYS task. Increases in the speech to pause ratio and significant decreases total pause time, mean pause length and percent pause were observed at 22 hours post-baseline.

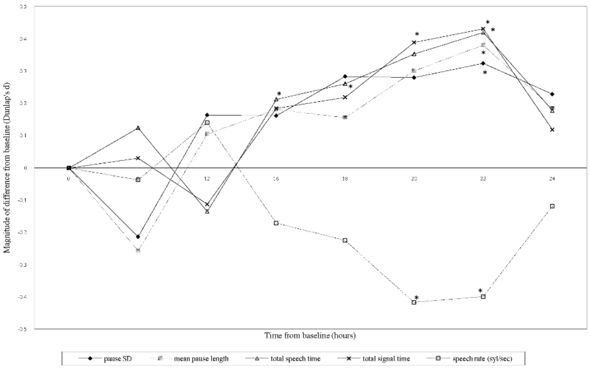

Figure 3: Magnitude of the difference (Dunlap’s d) in change from baseline performance of timing measures on the GRAN task over 24 hours.

3.3 Intensity measures

Increases in the alpha ratio were observed on the FREE task (Figure 4), with a peak at 22 hours, with a return closer to baseline during the final assessment. L_{eq} remained stable over the course of the experiment.

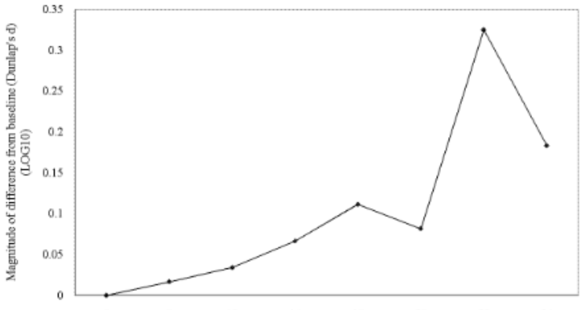

Figure 4: Magnitude of the difference (Dunlap’s d) in change from baseline performance of alpha ratio (LOG-transformed variable) on the FREE task over 24 hours.

4. Discussion

Altered acoustic output was observed as a consequence of increasing levels of fatigue induced by sustained wakefulness. Decreases in speaking rate and increases in mean pause length, signal time and total speech time were characteristic of participants’ performance on the reading task. The opposite was observed on the automated speech tasks. Speaking rate increased on the DAYS task and pause length decreased on the C120 task. Changes in acoustic output were also observed in data acquired during the AAAH task (F4 variation) and the FREE task (alpha ratio). The number of pauses did not change significantly from one recording to the next, nor did L_{eq}. These results are unique in the scientific literature as they show significant change over time in speech output due to relatively short periods of sustained wakefulness across different data types. Furthermore, the findings provide support for the use of voice acoustics in monitoring change resulting from fatigue in individuals over time.

4.1 Fundamental frequency

Clear changes in *f*0 variation (SD and CoV) were observed as a function of increasing levels of fatigue across the sustained vowel and all connected speech tasks. These findings are at odds with the qualitative judgments provided by Harrison and Horne [3] who reported reduced intonation (perceptual correlate of reduced *f*0 SD/CoV), after 36 hours of sustained wakefulness. *f*0 remained stable across the experiment contrary to previous work [5] which found changes in *f0* that were described to be “consistent with a circadian pattern” (p67). However, as intra-individual changes from baseline were not employed, it is difficult to determine the extent to which speech changed as a function of fatigue.

4.2 Formants

Formant patterns remained invariant despite increasing levels of fatigue, with the exception of F4 and F4 variation (SD/CoV). These three measures remained stable across repeated trials until 02:00 (18 hours sustained wakefulness) when marked changes were observed. A similar pattern of change was not found for F1-F3 (SD and CoV) ($d<0.2$). It is important to note that F4 variation has been shown to remain stable over a variety of inter-recording intervals (repeated trials, over a day and over a month) in healthy adults, suggesting that any change observed in this experiment was the result of fatigue rather than any systematic or technical error [6]. Although variation in F4 is not widely documented in the fatigue literature, some research has focused on mean F4 values or intensity levels around 3.5 kHz [15]. F4 has also been referred to as the ‘speakers formant’ [15] as the quality of an individual’s voice is thought to be determined, in part, by the steepness of the spectral slope.

4.3. Timing

Individual’s begun to show the affect of fatigue around midnight (16 hours post baseline), with maximal affect on timing arising just before dawn. Interestingly, the speech to pause ratio rose over the course of the experiment on the automated speech tasks, yet remained relatively stable on the reading task. Increases in the speech to pause ratio can arise from one of two reasons, either the total signal time increases while pause length remains the same, or pause length decreases while total signal length and number of pauses remains stable (as was the case for the automated speech tasks). The observation that participants produced automated speech tasks at the same rate throughout the experiment yet had greater speech to pause ratios as fatigue increased suggests that participants may employ a number of strategies to compensate for their increasing levels of fatigue (e.g., slowing rate speech rate while simultaneously lengthening the amount of time taken to produce each word). In doing so, pauses between words may be reduced by blending or co-articulating phonemes in adjacent words, thus concurrently increasing total speech time and restricting pause length, while maintaining total sample duration.

The multimodal processes described above were applied in a different context during the reading task. Significant increases in total sample duration, total speech time and mean pause length were observed as levels of fatigue were amplified. These processes combined manifest in a slower rate of speech (syllables per second). A reduced speaking rate may also be the result of an interplay between several factors including cognitive demand and fatigue induced changes to motor functioning (see Durmer and Dinges [16] for review).

4.3. Intensity

Clear increases in the alpha ratio were observed as a function of increasing levels of fatigue. These changes occurred in parallel with other acoustic measures of timing and frequency. Measures of spectral tilt have not previously been considered in the sustained wakefulness literature, however, studies focusing on vocal fatigue point to the importance of spectral energy in monitoring vocal change [17]. Prolonged vocal use (vocal loading) has been linked to changes in vocal stability [18] and several acoustic correlates including *f0*, spectral tilt (alpha ratio) and formant frequencies [19] appear sensitive to changes induced through such experimental models. As discussed, increases in the alpha ratio are often considered a reflection of increases in vocal adduction [20], and also a reflection of changing levels of vocal loudness [13]. However, as L_{eq} remained stable over the course of the experiment, a link between vocal loudness and spectral balance was not clearly established.

5. Conclusion

Previous studies examining changes in psychomotor performance over twenty-four hours of sustained wakefulness [21], revealed a steady increase in reaction time from 16 hours post baseline assessment onwards with a return closer to baseline performance during the last assessment. The pattern of decline and renewal of psychomotor performance in the cognitive assessments is mirrored by the speech changes observed in the current study. This evidence provides weight to the argument that observed changes in speech were in fact the result of sustained wakefulness, and not the effect of systematic or random error resulting from repeated exposure to the stimuli. Given these data, acoustic analysis of speech appears to provide one means of objectively monitoring intra-individual change in adults without vocal pathology.

6. References

[1] D. M. Gaba and S. K. Howard, "Fatigue among Clinicians and the Safety of Patients," *New England Journal of Medicine,* vol. 347, pp. 1249-1255, October 17, 2002 2002.

[2] E. G. Bard, *et al.*, "The DCIEM Map Task Corpus: Spontaneous dialogue under sleep deprivation and drug treatment," *Speech Communication,* vol. 20, pp. 71-84, 1996.

[3] Y. Harrison and J. A. Horne, "Sleep deprivation affects speech," *Sleep,* vol. 20, pp. 871-7, Oct 1997.

[4] H. P. Greeley, *et al.*, "Fatigue estimation using voice analysis," *Behavior Research Methods,* vol. 39, pp. 610-619, August 2007 2007.

[5] J. Whitmore and S. E. Fisher, "Speech during sustained operations," *Speech Communication,* vol. 20, pp. 55-70, 1996.

[6] A. P. Vogel, *et al.*, "Reliability, stability and sensitivity to change and impairment in measures of timing and frequency," *Journal of Voice,* in press.

[7] T. Roehrs and T. Roth, "Multiple Sleep Latency Test: technical aspects and normal values," *Journal of Clinical Neurophysiology,* vol. 9, pp. 63-7, Jan 1992.

[8] P. Boersma and D. Weenink, "Praat: doing phonetics by computer ", 5.0.46 ed, 2009, p. Computer Program.

[9] A. P. Vogel, *et al.*, "Standardization of pitch range settings in voice acoustic analysis," *Behavior Research Methods,* vol. 41, pp. 318-324, 2009.

[10] K. M. Rosen, *et al.*, "Automatic method of pause measurement for normal and dysarthric speech," *Clinical Linguistics & Phonetics,* vol. 24, pp. 141-154, 2010.

[11] A. M. Laukkanen, *et al.*, "Changes in voice and subjective sensations during a 45-min vocal loading test in female subjects with vocal training," *Folia Phoniatrica et Logopaedia,* vol. 56, pp. 335-46, 2004 2004.

[12] M. Nordenberg and J. Sundberg, "Effect on LTAS of vocal loudness variation," *Logopedics Phoniatrics Vocology,* vol. 29, pp. 183-91, 2004.

[13] J. Sundberg and M. Nordenberg, "Effects of vocal loudness variation on spectrum balance as reflected by the alpha measure of long-term-average spectra of speech," *Journal of the Acoustical Society of America,* vol. 120, pp. 453-7, Jul 2006.

[14] W. P. Dunlap, *et al.*, "Meta-analysis of experiments with matched groups or repeated measures designs," *Psychological Methods,* vol. 1, pp. 170-177, 1996.

[15] I. V. Bele, "The speaker's formant," *Journal of Voice,* vol. 20, pp. 555-578, 2006.

[16] J. S. Durmer and D. F. Dinges, "Neurocognitive Consequences of Sleep Deprivation," *Seminars in Neurology,* vol. 25, pp. 117-129, 2005.

[17] L. Rantala, *et al.*, "Working-Day Effects on the Spectral Characteristics of Teaching Voice," *Folia Phoniatrica et Logopaedia,* vol. 50, pp. 205-211, 1998.

[18] A. M. Laukkanen, *et al.*, "Acoustic measures and self-reports of vocal fatigue by female teachers," *Journal of Voice,* vol. 22, pp. 283-9, May 2008.

[19] V. Jonsdottir, *et al.*, "Changes in teachers' voice quality during a working day with and without electric sound amplification," *Folia Phoniatrica et Logopaedia,* vol. 55, pp. 267-80, Sep-Oct 2003.

[20] J. Gauffin and J. Sundberg, "Data on the glottal voice source behavior in vowel production," *Speech Transmission Lab. Q. Program Status Report,* vol. 21, pp. 61-70, 1980.

[21] P. Maruff, *et al.*, "Fatigue-related impairment in the speed, accuracy and variability of psychomotor performance: comparison with blood alcohol levels," *Journal of Sleep Research,* vol. 14, pp. 21-27, 03 2005.