Acoustics'08
Paris
June 29-July 4, 2008
www.acoustics08-paris.org

Analysing cockpit and laboratory recordings to determine fatigue levels in pilots' voices

Robert Ruiz[a], Philippe Plantin De Hugues[b] and Claude Legros[c]

[a]Université Toulouse II, LARA, 5 allée Antonio Machado, 31058 Toulouse Cedex, France
[b]Ministère de l'Ecologie, de l'Energie, du Développement Durable et de l'Aménagement du Territoire, 200 rue de Paris, Bât 153, Aéroport du Bourget, 93352 Le Bourget Cedex, France
[c]Université de Toulouse-Le Mirail, Laboratoire d'Acoustique LAUTM, 5, allées Antonio Machado, 31058 Toulouse Cedex 1, France
legros@univ-tlse2.fr

Analysis of pilots' voices was undertaken under controlled laboratory conditions, using aeronautical terminology, standard professional equipment and a standardized recording environment. The aim of the experiment was to determine the ways in which the acoustic characteristics of the voice are modified after a phase of sleep. The results, which are presented in this paper, show very significant variations in the dispersion parameters, in cases such as jitters, associated with the fundamental frequency.
In an aircraft cockpit, the electro-acoustic environment is noisy. Therefore, it is essential to improve and adapt voice analysis methods to achieve reliable results. Cockpit Voice Recorder recordings have poor acoustic characteristics due to sound acquisition system quality and the signal-to-noise ratio, in addition to the non-predetermined vocabulary range.
This paper will outline the modifications made to the laboratory methodology and the comparative results obtained, to improve the analysis of CVR recordings.
Comparisons of the test recordings meant that it was feasible to determine, via voice analysis, the state of drowsiness of a pilot and, made possible the study of CVR recordings from accidents using these new analytical techniques.

1 Introduction

Many studies have shown the existence of a link between the voice and the emotions of a speaker [1, 7, 8]. The situations which cause emotional disturbance are many and varied, such as psychomotor exercises [2], observation of images [3], real or simulated aircraft accidents [1].
In the aeronautical context, emotions in the voice are sought on Cockpit Voice Recorders (C.V.R) and on recordings made in simulated conditions.
The general purpose of this study is to proceed with the same approach for pilot's fatigue.
Firstly, it aims to confirm that voice changes may occur when a speaker is tired. This is the purpose of the experimentation on sleep inertia described below (paragraph 2). Secondly, voice analysis laboratory methods have to be adapted to Cockpit Voice Recording analysis (paragraph 3).
Based on this set of results, it will be possible to study fatigue in real flight conditions both from the C.V.R and from the analysis of recordings made in the aircraft cockpit.

2 Voice and sleep inertia: laboratory experiment [4]

Two factors must be taken into account for the study of workload-induced fatigue. The first one is fatigue accumulated during periods of activity. The second one is the effects of drowsiness which are presented here.
Pilots who are subject to daily short-haul rotations can be in a state of drowsiness late in the day or even during early morning flights. The question that arises is whether the ability to concentrate and work efficiency are always in the optimum level.
Medical measurements have been made and are completed by a collection of non-invasive voice data.

2.1 Experimental conditions

The experiments were conducted with a pilot in a French hospital (sleep disorders department).
ElectroEncephaloGrams and ElectroCardioGrams were carried out by a specialized laboratory. The voice was recorded on D.A.T by way of a headset microphone ensuring a constant distance from the mouth.
Three recordings were made:

Recording 1: upon arrival in the lab (11 AM)
Recording 2: after lunch (2 PM)
Recording 3: The pilot was taken to a room where his sleep was monitored. A few minutes after falling asleep he was suddenly awakened by a very powerful light. He then had to perform a number of tasks on a computer similar to those usually performed in flight (3 PM).

The pilot reads the same five sentences in each of the three recordings. They came from aeronautical terminology. A sample sentence was : « Bravo, Victor, Charlie montez au niveau deux cinq zéro » (« Bravo, Victor, Charlie climb level two five zero »).

2.2 Voice analysis

108 vowels were segmented by selecting the time period where the signal is quasi-stationar: 37 vowels for recording 1, 30 for recording 2 and 41 for recording 3.
The analyses were performed using Matlab laboratory programs. The parameters calculated for each vowel were: the mean fundamental frequency $<F_0>$, the associated standard deviation (σ), the Variation Coefficient (V.C), the mean jitter (M.J), the jitter factor (J.F), the shimmer (S).

$$V.C = 100 \times \frac{\sigma}{< F_0 >} \text{ (in \%)} \tag{1}$$

$$M.J = \frac{1}{n-1}\sum_{i=1}^{n-1}\left|F_0(i) - F_0(i+1)\right| \text{ (in Hz)} \tag{2}$$

$$J.F = 100 \times \frac{M.J}{< F_0 >} \text{ (in \%)} \tag{3}$$

$$S = \frac{1}{n-1}\sum_{i=1}^{n-1}\left|20.\log\left(\frac{A_i}{A_{i+1}}\right)\right| \text{ (in dB)} \tag{4}$$

n is the number of pitch periods of the segmented vowel wave form and A_i is the amplitude of the first major peak of the i period. Significant results were obtained for all these parameters except for the mean fundamental frequency and the shimmer (Tables 1 and 2). But as shown in Figures 1 and 2 the only visual examination indicates

high increases of values at the beginning of recording 3. The same observation is made for standard deviation, mean jitter and jitter ratio For shimmer the visual examination shows a tendancy (Figure 6) but the increase is not significant (Table 2).
An example of a vowel phase portrait issued from recording 3 also shows the important cycle-to-cycle variations of the fundamental frequency (Figure 5) comparing the scale with a phase portrait during recording 2 (Figure 4).

	Recording 1+2 (67vowels)	Recording 3 (41 vowels)
$\langle F_0 \rangle$ (in Hz)	129,92 (21,70)	124,41 (24,58)
σ (in Hz)	21,86 (17,87)	57,39 (24,29)
V.C (in %)	16,80 (11,92)	47,27 (21,04)
M.J (in Hz)	26,67 (17,67)	78,17 (28,90)
J.F (in %)	20,86 (13,44)	62,97 (23,02)
S (in dB)	2,68 (1,74)	4,84 (2,10)

Table 1 Mean values (standard deviation into brackets) of vowel features between neutral period (recording 1+2) and « stress » one (recording 3).

	F-test (significance level 0,01)			t-test (significance level 0,01)		
	F	$F_{critical}$	p	t	$t_{critical}$	p
$\langle F_0 \rangle$	1,28	2,04	0,366	-1,22	2,36	0,113
σ	1,85	2,04	0,026	8,73	2,36	2.10^{-14}
V.C	3,11	2,04	4.10^{-5}	8,48	2,39	6.10^{-12}
M.J	2,67	2,04	4.10^{-4}	10,3	2,39	4.10^{-15}
J.F	2,93	2,04	10^{-6}	10,65	2,39	2.10^{-15}
S	0,69	0,48	0,09	-5,43	2,65	4.10^{-7}

Table 2 Fisher-Snedecor and Student tests between neutral period (recording 1+2) and « stress » one (recording 3).

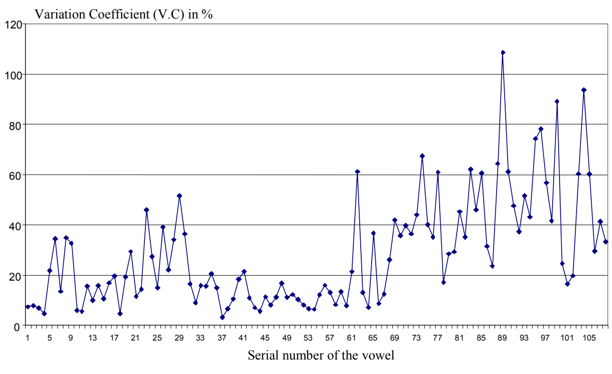

Fig.1. Variation Coefficient of the Fundamental Frequency versus serial number of the vowel.

Starting from the 67th analysed vowel, which corresponds to waking the speaker, the associated parameters of the fundamental frequency varied significantly (Table 2). It should be noted that the mean fundamental frequency did not present any visible and significant variation for recording 3 (Figure 3, Table 2).

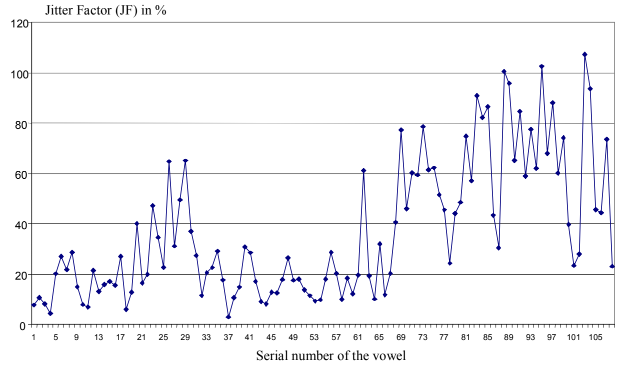

Fig.2. Jitter Factor of the Fundamental Frequency versus serial number of the vowel.

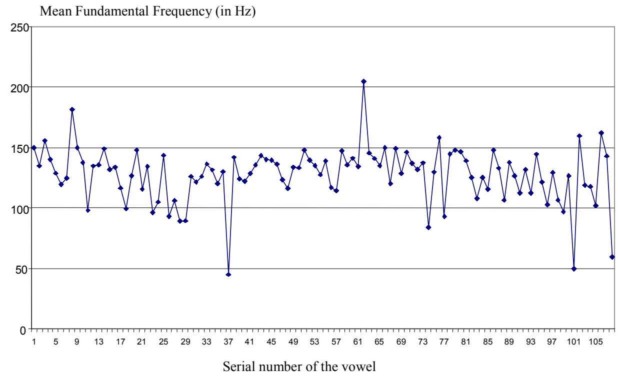

Fig.3. Fundamental Frequency versus serial number of the vowel.

Sleep inertia can modify a pilot's voice: an increase of short-term instability of the fundamental frequency is then observed. The dispersion parameters (standard deviation, variation coefficient, jitters) vary significantly between the ordinary periods of voicing (recordings 1 and 2) and the « stress » one (recording 3).

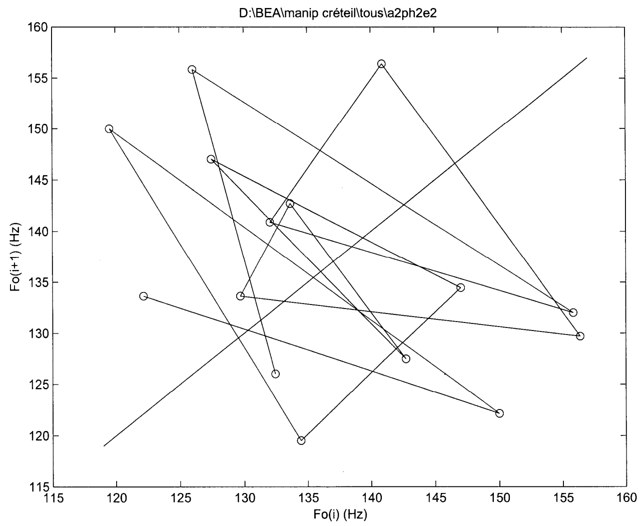

Fig 4. Phase portrait of a vowel issued from recording 2.

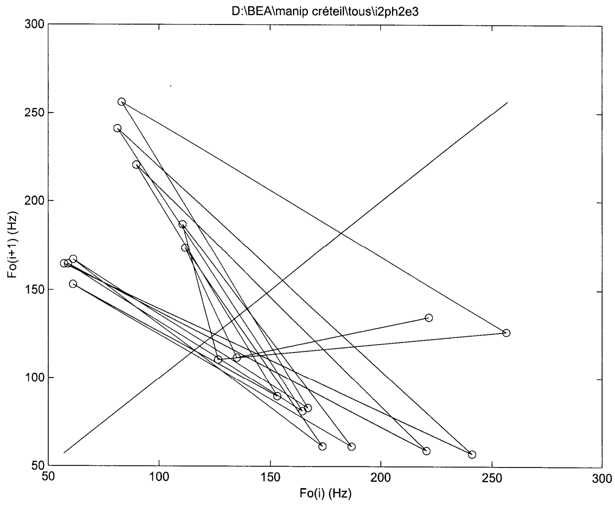

Fig.5. Phase portrait of a vowel issued from recording 3

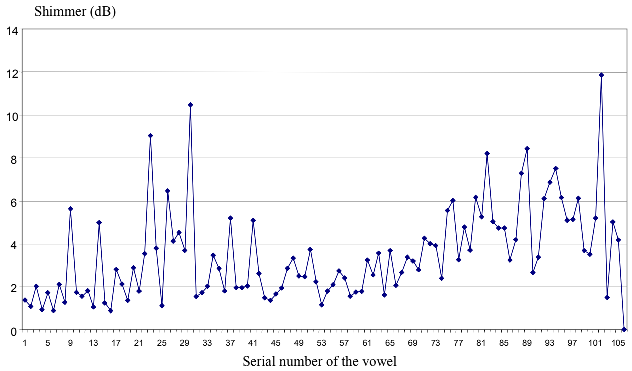

Fig.6. Shimmer versus serial number of the vowel

3 Analytical methods and Cockpit Voice Recordings [5]

3.1 Bandwidth

The low-frequency limit of the C-V-R bandwidth (150 Hz) prevents detection of the fundamental frequency after low-pass filtering. Here, the approach is based on the detection of peak amplitudes on the vowel signal. This makes the measurement independent of the system's low cut-off frequency.

The mean fundamental frequency remains almost identical to that of the cepstral analysis.

3.2 Signal-to-Noise Ratio

This parameter was high in laboratory conditions. At equivalent speech signal quality, the level of background noise on the C-V-R tape reduces the signal-to-noise ratio. In most cases there is no impact on the estimation of parameters related to the fundamental frequency, but spectral estimation is disrupted.

Spectral analysis of the background noise taken from a C-V-R shows relatively constant levels for frequencies between 1000 Hz and 3000 Hz and a significant power level of around 500 Hz (Fig. 7). The spectral slope is about +5dB/octave up to the 2000 Hz/octave.

The sound level of background noise and each vowel are added into the recorded signal and then the estimation of the vowel spectrum requires specific treatment. Indeed, the natural decrease of vowel sound level is hidden by the noise spectral slope and the peak at 500 Hz may lead to an overestimation of the sound level around the first formant.

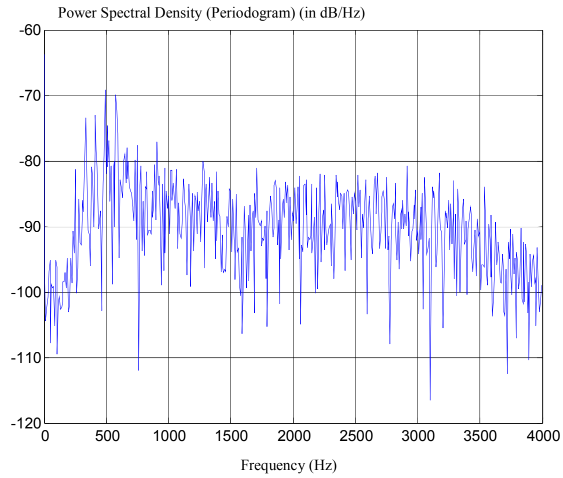

Fig.7. Background Noise Spectrum.

3.3 Modified spectral estimation

The spectral analysis of vowels is performed by an all-pole spectral model obtained from linear prediction (LPC autocorrelation method) [6].

In laboratory conditions pre-emphasis (simple one-zero filter of the form $1-\mu.z^{-1}$ with μ near or equal to one) is applied to estimate the transfer function of the vocal tract without taking into account the effects of the lips radiation and those of the glottal wave. A lack of pre-emphasis ($\mu = 0$) leads to an estimate of the spectrum of the vowel.

With the presence of C-V-R background noise, pre-emphasis enhances the spectral effect of noise for high frequencies (Figure 8).

The slope of the frequency response of the vocal tract shows an increase of roughly +11 dB/octave in addition to its normal slope: +6 dB/octave due to pre-emphasis and +5dB/octave due to noise.

Then for the spectrum vowel estimation the increase is about +5 dB/octave because pre-emphasis is not used. Nevertheless, sprectrum shows excessive high frequency levels due to the presence of background noise in the recorded signal (Figure 9).

Consequently, the best way is to proceed without pre-emphasis and to perform a first order low-pass filtering before the LPC analysis (Figure 10). The choice of the cutoff frequency is related to the background noise level at 500 Hz.

The analysis conditions are: order 12 for the 8000 Hz sampling frequency, 512 samples in the Hamming window analysis, recovery 50%, pre-emphasis coefficient μ equal to 0.98 or 0 (Figures 8, 9 and 10).

The noise spectral slope effect on the recorded signal is eliminated to give an estimation of the vowel spectrum.

The case of the sound level peak around the 500 Hz frequency is not attributable to the pilot's voice.

The background noise analysis shows an increase of sound level slightly less than 20 dB at 500 Hz compared to the 1300 Hz - 3300 Hz frequency band (Figure 7). The low-pass first order filtering attenuation is of 18dB at 500 Hz from the 63 Hz cutoff frequency on the recorded signal level (vowel + background noise).

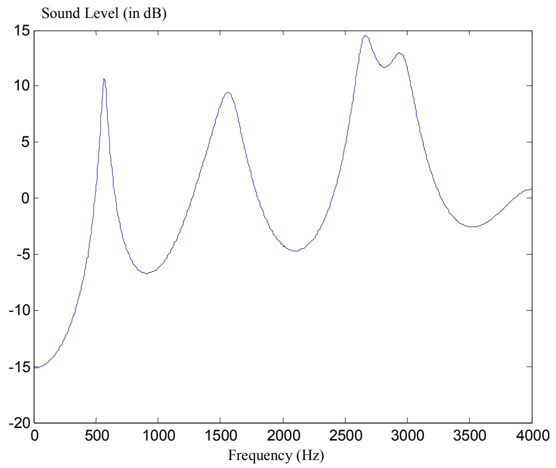

Fig.8. Vowel L.P.C spectrum (vowel + background noise) with pre-emphasis (μ=0.98).

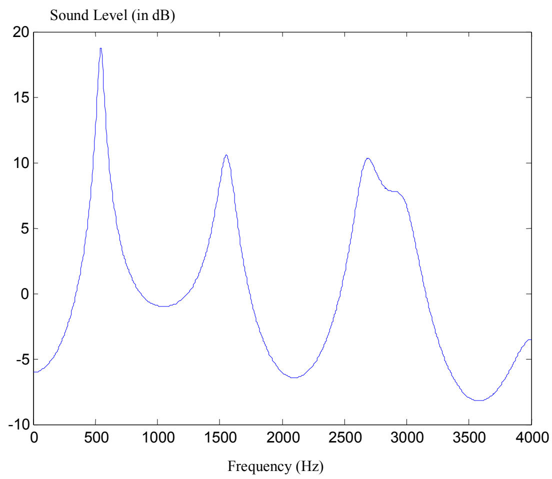

Fig.9. Vowel L.P.C spectrum (vowel + background noise) without pre-emphasis (μ=0).

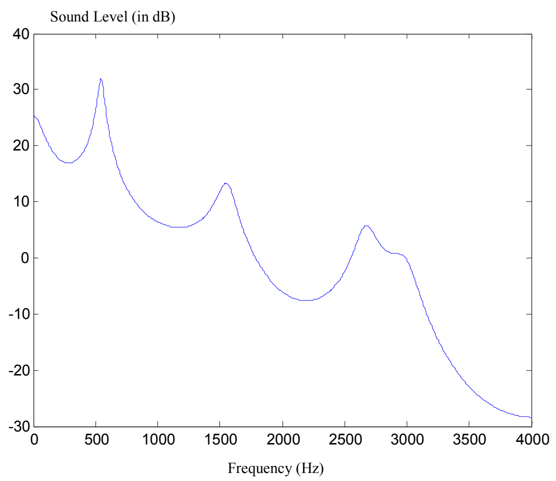

Fig.10. Vowel LPC spectrum (vowel + background noise) without pre-emphasis and with low-pass filtering before sprectrum calculation

The low-pass filtering reduces the excessive level at 500 Hz. However, it is not an accurate correction because the spectral shape of reinforcing around 500 Hz is not taken into account by filtering.

The correction is nevertheless effective enough to provide a suitable estimation of the vowel spectrum (Figure 10).

Conclusion

Even if only one pilot's voice has been analysed, and if the background noise can be modified in other flying situations, the set of results presented in this paper encourages development of new experiments in voice analysis in real conditions.

The sleep inertia experiment indicates that pilot's voice recordings in the cockpit can provide, at the least, significant variations of fundamental frequency dispersion parameters.

Analysis of voice recordings on C-V-R's will also be possible with the spectral corrections demonstrated above.

References

[1] R.Ruiz, C.Legros, A.Guell, "Voice Analysis: Application to the Study of the Influence of a Workload," *J. Acoustique*, vol. 3 (2), pp. 153-159 (1990).

[2] R.Ruiz, E.Absil, B.Harmegnies, C.Legros, D.Poch, "Time- and Spectrum-Related Variabilities in Stressed Speech under Laboratory and Real Conditions," *Speech Com.*, vol. 20 (1-2), pp. 111-129 (1996).

[3] R.Ruiz, R.Da Silva Neves, C.Martinot, S.Vautier, "Interactions between Emotional Content of Pictures and Acoustic Features of Speech", *17th International Congress on Acoustics,* Roma, Italy, 6A 14.04 pp 276-277 (2001).

[4] C.Legros, R.Ruiz, "Etude de l'influence d'une période de sommeil sur la voix de pilotes d'avion ", *rapport final de l'étude BEA N°3304/2005*, (Mars 2007).

[5] R.Ruiz, "Prise en compte des spécificités électroacoustiques des communications avion – tour de contrôle pour l'adaptation des mesures de laboratoire sur le signal vocal, phase 2", *rapport final de l'étude BEA N°3841/2007,* (Novembre 2007).

[6] J.D.Markel, A.H.Jr.Gray, Linear prediction of speech, Springer-Verlag, (1982).

[7] "Speech under stress", p*roceedings of the E.S.C.A-N.A.T.O tutorial and research workshop on Speech under Stress*, Lisbon, Portugal, (1995).

[8] J.H.L. Hansen, C.Swail, A.J.South, R.K.Moore, H.Steeneken, E.J.Cupples, T.Anderson, C.R.A.Vloeberghs, I.Trancoso, P.Verlinde, "The impact of speech under stress on military speech technology*", NATO Project Report AC/232/IST/TG-01*, (March 2000).