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

Spectrum-related variabilities in stressed speech under laboratory and real conditions

Robert RUIZ (1), Emmanuelle ABSIL (2), Bruno GRAMATICA (1), Bernard HARMEGNIES (2), Claude LEGROS (1), Dolors POCH (3).

(1) Laboratoire d'Acoustique, Université Toulouse-Le Mirail
5 allées Antonio Machado, 31058 Toulouse Cedex, France.

(2) Laboratoire de Phonétique, Département de communication parlée,
Université de Mons-Hainaut, 18 place du parc, B-7000 Mons, Belgium.

(3) Laboratori de Fonetica, Departament de filologia espanyola,
Facultat Lletres, Universitat Autonoma de Barcelona, 08193 Bellaterra, Spain.

ABSTRACT

Emotional stress induced by various types of situations leads to vocal signal modifications. Previous studies have indicated that stressed speech has a higher fundamental frequency Fo and presents noticeable changes in the spectrum of vowels.

In this paper, spectrum-related variations are studied. All the corresponding spectrum-related indicators are applied to stress detection in a laboratory situation (Stroop's psychometric stress test) and a real situation (Cockpit Voice Recording) inducing emotional disturbances.

1. INTRODUCTION

Stress induces vowel spectral modifications. Comparative observations of sonagrams at rest and under stress (or simulated emotions) have indicated changes in shape (e.g attack, regularity of successive pulses, accuracy of articulation...) [1], and modification of the amount of high frequency energy [1,6]. Such effects are difficult to quantify and sonagrams are very useful to show that. The analysis of octave and third octave long term average spectra came to the same conclusion [6,7]. A specific indicator from third octave spectral analysis has been introduced to evaluate stress in different phases of a Russian space flight. It also indicates an increasing high frequency energy with an increasing heart rate [5].

Studies of formant variations under stress have shown interesting results. In [7], the authors have noticed modifications of the uniform formant structure, and shifts in the direction of increasing two first order formant frequencies are observed under stress induced by a simulated motion disorientation task [2].

Finally, it seems that spectral modifications occur when a speaker is in a state of stress. But this state is very often induced by different stressing tasks or situations, and the reliability of the spectrum-related indicators is not evaluated both in laboratory and real conditions. The purpose of this paper is to test spectral characteristics for a Cockpit Voice Recording of two pilot's voices during a severe flight incident and for a recording of a subject undergoing a Stroop's psychometric stress test.

2. METHOD

Two major kinds of modifications exist: energy displacements and shifts of formants frequencies.

• The first ones are evaluated by the estimation of Cumulative Spectral Probability Diagrams (CSPD) in two complementary bands of the spectrum [3]. More than a simple comparison of energy this measure indicates the way sound level is distributed. Two different spectra can have the same energy in two bands but not necessarily the same CSPD [4]. CSPD-related characteristics studied are the two following:

- Δ is the area between low and high frequency CSPD with a frequency limit equal to the second antiresonance of each spectrum;

- Fs is the spectral balance frequency: i.e frequency for which Δ is minimal.

Envelope spectra result from an LPC model of each frame of the signal. A mean spectral envelope of the steady-state vowel is obtained by averaging the sets of prediction coefficients. CSPD are obtained with a 1dB class width and smoothed by polynomial regression. Signal is low-pass filtered at 5KHz, sampled at 10KHz with a 75% overlapping. The linear prediction model of each pre-emphasized and Hamming weighted frame is of order 15.

• For each vowel, we evaluated its first, second ant third formants at its center, by means of a KAY DSP 5500 analyser, on the basis of wide band spectrograms, FFT spectra and LPC formant tracking routine.

Formants values in Hz have been converted into Mels prior to statistical treatment. Two-way analyses of variance have been performed in order to test for the existence of stress-related formants variabilities, notwithstanding the effect the vowel. Three analyses of this kind were performed : one for each formant. In order to locate the effects, one-way analyses were also performed vowel by vowel, in case a significant stress effect had been observed in the two-way analysis.

Only tests for which the null hypothesis probability is little ($p < .1$), either for a stress effect or for a stress-vowel interaction effect, are reported in this paper. The previous analysis of variance required to consider each formant in turn. Such analyses are interesting, since they allow microscopic study of the effects of stress. Nevertheless, they do not provide information about the global dynamics of the reported effects. A macroscopic viewpoint can be adopted by considering each vocalic sound uttered by the speaker as a point in a multidimensional space (in this case, the F1-F2-F3 space). Than the question of whether these vowel-points form clusters with variable degrees of consistency may be addressed. One way of doing this is to test whether the distances to the center of the formants space depend upon stress. This requires to choose a reference center for the formants space. Schwa, a theoretical vowel with formants values regularly spaced in the frequency domain (500 Hz, 1500 Hz, 2500 Hz, ...) is often regarded as such. In order to cope with inter speaker variabilities, it is also possible to evaluate each subject's own F1-F2-F3 space center.

3. EXPERIMENTAL CONDITIONS

3.1 Real situation stress

The Cokpit Voice Recorder of a crashed aircraft is studied. The timing of the events during the last thirty minutes shows the occurence of two successive failures in flight commands. These are followed by discussions between the pilots about the reality and causes of the events. Finally, after a third occurence, the plane crashes. It is obvious that this record contains a stressful phase (at the end of the flight) and several other phases with lower stress (during discussions after the appearance of technical problems). A rest period exists before the flight incident. Therefore, three possible degrees of emotional behaviour are in evidence: a severe one at the end of the flight, a less important one for the discussions and the previous occurences, and a rest one. They will be referred to as levels 2,1 and 0 of the independent variable stress.

3.2 Laboratory situation stress

The stressor consisted of a French version of the so called "Colour-Word Test" by Stroop (1935). The test is divided into three different phases (referred to as phases 1, 2 and 3), within each of which the subject has to utter colour names as quickly as he can. The first two phases may be considered as control situations, although the third is the one giving rise to stress. During the first phase, the subject has to read aloud 50 groups of 4 colour names (in this case: "vert", "jaune", "rouge", "bleu"). The words are written in black on white sheets, presenting 50 rows, each containing the 4 colours'names organised in random order. For the second phase, the subject is presented 50 rows of 4 coloured squares each. The colours are again green, yellow, red and blue. The subject has to tell the colour of each square. For the third part of the test, the subject is given 50 rows of 4 colours names each. This time, the words are nevertheless not written in black: the ink used for each word is of a different colour than the one refered by the word itself. The subject has to tell what is each word's ink colour (e.g., the word "yellow"is written in blue and the subject has therefore to say "blue"). /u/ of "rouge", /ɛ/ of "vert", /o/ of "jaune", /ɸ/ of "bleu" are the selected vowels of the French speaker which are analysed in the three phases.

4. RESULTS

4.1 Frequency distribution of sound level

Steady-state French vowels are extracted and analysed all along the recording: /a/ for the CVR and /u/ /ɛ/, /o/, /ɸ/ for the Stroop test.

4.1.1 Δ area

Both for real and laboratory situations, variations of Δ do not indicate a unique sense of variation with time, i.e with increasing stress degree. Low and high values of Δ exist, indicating motions of sound level distribution around the second antiresonance frequency. Even if very low values are measured in the real situation, it is not possible to correlate them with a stress emergence. The changes occur all along the recording. In the case of the laboratory situation lower Δ are measured at the beginning of the third phase of the stress test only for the colour "vert" (continuous line figure 1) and "bleu" (dotted line figure 1).

4.1.2 Spectral balance frequency Fs

The major result concerning the spectral balance frequency Fs is observed when plotting it versus another characteristic, for example the fundamental frequency Fo of the same vowels (figure 2). Similar plots are obtained for the pilot and for the speaker of the laboratory stress test (figure 3). For each type of vowel studied /a/, /ɛ/, /ɸ/, /u/, /o/, even if the spectrum form is different, three classes of Fs can be defined: two groups containing extrema Fs values (lower and higher ones) and a class (sometimes several) with intermediary ones.

Direct conclusions linking these groups with degrees of stress are not possible for the moment. For the real situation, presumed stressed vowels belong to all groups. But for the laboratory test one can notice that a majority of third phase vowels belong to center groups of the plot. This result suggests that a relation between spectral balance frequency and stressed speech is conceivable.

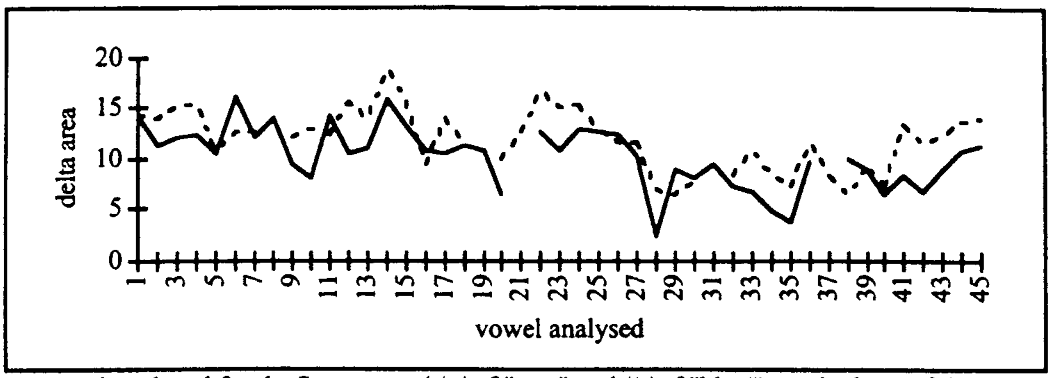

Fig.1: Δ versus vowel analysed for the Stroop test (/ε/ of "vert" and /ɸ/ of "bleu"); each phase of the test has 15 vowels.

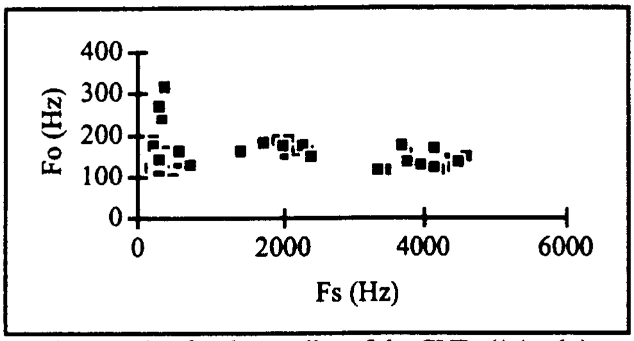

Fig.2: Fo/Fs for the copilot of the CVR. (/a/ only)

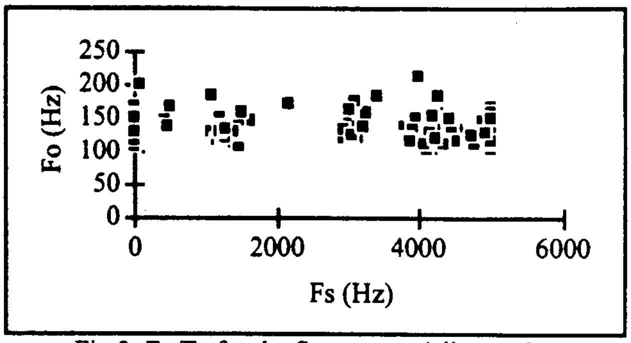

Fig.3: Fo/Fs for the Stroop test (all vowels).

But now a systematic study of Fs is necessary in order to conclude about the particular, or not, nature of these measures.

4.2 Formants frequencies

For the laboratory stress situation, formants analyses were performed on a subsample of 60 vowels in each phase (out of 200 by phase), allowing context invariance for the selected vowels. For the real situation stress, the analyses were performed on a restricted, balanced sample of words extracted from the three periods of the flight. The aim was to cope with the possible important effect of vowel phonetic environment. For that purpose, only vowels with the same anterior context in the three subsamples were taken into account : the vowel is selected only if consonant-vowel combination appears in all three parts of the CVR corpus.

4.2.1 Real situation stress

Results of the treatments are summarized in Table1:

Pilot	Copilot
F2 (p=.002) /a/↗ /e/↗ /ε/↗	F3 (p=.005) /u/↘
F3 (p=.01) /ε/↗ /ã/↗	F1 (p=.08) /i/↗ /u/↘
F2 (p=.03) stress-vowel interaction	F1 (p=.08) stress-vowel interaction

Table1: probabilities under H_0 (p) for the significant tests (arrows indicate the sense of the frequency variation under the effect of stress).

For the pilot, we note a significant effect of stress on the second formant. Detailed analyses reveal frequencies increases in vowels /a/, /e/, and /ε/. A significant variation of F3 with formant frequencies increases for /ε/ and /ã/ is also to be noted. Finally, a significant interaction effect appears between stress and vowels variables in F2.

For the copilot, the analysis reveals a significant effect of variable stress on F3. In this case, third formant values of /u/ decrease. We also observe an effect on the first formant, consisting in frequency increase for /i/ and frequency decrease for /u/ values. A significant interaction is also observed between stress and vowel, for the first formant.

A similar treatment, performed on data from stress levels 0 and 1, shows significant effects of stress on F2, for the pilot (p=.02), and on F3 for the copilot (p=.046).

4.2.2 Laboratory situation stress

As in the CVR data treatments, formants values have been converted into Mels, prior to the analyses of variance.

The results are presented in Table2.

Considering all the phases of the experiment, significant differences appear in each formant. The first one is characterized by variations in /u/ : the mean value increases between phases 1 and 2, then decreases between phases 2 and 3. First formant frequencies decrease for /ɸ/ and increase for /ε/ under the effect of stress. Variations in F2 are important in vowel /ε/.

F1	F2	F3
Phase ($p=.016$) /u/↗ ↘ /ϕ/↘↘ /ε/↗↗	Phase ($p<.001$) /ε/↗↗	Phase ($p=.083$) /u/↘↘ /ϕ/↘ ↗ /ε/↗↗
Phase-vowel interaction ($p=.001$)	Phase-vowel interaction ($p<.001$)	Phase-vowel interaction ($p<.001$)

Table2: probabilities of H_0 for the significant tests ; the arrows indicate the sense of variation of the involved formant, from phase 1 to phase 2 (left arrow), and from phase 2 to phase 3 (right arrow).

Finally, F3 variations are observed in some vowels : decrease for /u/, increase for /ε/ and increase after decrease for /ϕ/. Situation-vowel interaction is highly significant for each formant. The treatment performed on data from situation 1 and 2 only shows very significant between-phase differences for the second formant. We observe in this case an increase of the mean values for the vowels /u/ and /ε/.

4.3 Distances to the F1-F2-F3 space center

4.3.1 Real situation stress

Values of the three first formants of vowels /oe/ and /ə/ may be used for determining F1-F2-F3 space center (Table3).

Speaker	F1	F2	F3
Pilot	605	1 149	1 626
Copilot	524	1 209	1 584

Table3: central frequencies: mean values (in Mels)

Data processing consists in 2-way analyses of variance, with stress and vowel as independent variables, and the Euclidean distances of the analysed sounds to the F1-F2-F3 space center as dependent variable.

Independent variable	Pilot	Copilot
Stress	$p = .021$	$p = .556$
Interaction	$p = .539$	$p = .003$

Table4: probabilities of H_0 for the analysis of variance.

For the pilot, results indicate an influence of the more stressing situation of the flight on the *distance to center* parameter. No stress-vowels interaction is noticed here. A reversed configuration can be observed in the copilot's data (no significant effect of stress, but a significant stress-vowel interaction).

4.3.2 Laboratory situation stress

The same treatments as those performed on CVR data have been applied on Stroop test corpus. Nevertheless, since the recordings contained tokens of the French vowels /u/, /ϕ/, /ε/ and /o/ only, the theoretical Schwa vowel had to be considered the center of the F1-F2-F3 space. No significant effect of the phase is revealed by the analysis of variance ($p=.41$). The phase-vowel interaction effect is however very significant ($p<.001$), which can be interpreted in terms of effects in specific vowels under the Stroop situation.

5. CONCLUSION

Results have shown two types of spectral modifications induced by stress: changes in sound level distribution and shifts of formants. But variations are not systematic neither for the real situation nor for the laboratory one. A stressing factor (incident, third phase of the Stroop test) does not necessarily imply an increase or a decrease of the spectrum-related characteristics. Variations are not closely linked with stressor appearance. Nevertheless, results like the decrease of Δ, the groups of Fs and the shifts of formant frequency occur for different steady-state vowels. This indicates that modifications are closely connected with vocal expression of stress.

6. REFERENCES

[1] HECKER M.H.L., STEVENS K.N, Von BISMARK G., WILLIAMS C.E., (1968) "Manifestations of task induced stress in the acoustic speech signal", J. Acoust. Soc. Am., 44 n°4, 993-1001.

[2] MOSKO J. D., STEVENS K. N., GRIFFIN G. R. (1983) "Interactive voice technology: variations in the vocal utterances of speakers performing a stress-inducing task", Naval Aerospace Medical Research Laboratory Report 1300, Pensacola FA, USA.

[3] RUIZ R., LEGROS C. (1994) "The cumulative spectral probability diagram: theory and experiments", Acta Acustica, 4, 215-222.

[4] RUIZ R., LEGROS C. (1995) "Vowel spectral characteristics to detect vocal stress", 15th Int. Congress on Acoustics, Trondheim, Norway.

[5] SIMONOV P. V., FROLOV M. V., "Utilization of human voice for estimation of man's emotional stress and state of attention", Aerospace Med., 44(3), 256-258.

[6] WILLIAMS C. E., STEVENS K. N. (1981) "Vocal correlates of emotional states", In Speech evaluation in psychiatry, edited by J.K Darby, Jr., by Grune & Stratton, Inc., 1238-1250.

[7] WILLIAMS C. E., STEVENS K. N. (1972) "Emotions and speech: some acoustical correlates", J. Acoust. Soc. Am., 52, 1238-1250.

ACKNOWLEDGEMENTS

The "Services Techniques de la Navigation Aérienne" provided the magnetic recording.