ACOUSTICS 2012

Amplitude Modulation of Vowel Glottal Pulses - Application to Sleep Inertia

R. Ruiz[a], P. Plantin De Hugues[b] and C. Legros[c]

[a]LA.boratoire de Recherche en Audiovisuel (LA.R.A), Université Toulouse-le Mirail, 5 allées Antonio Machado, 31058 TOULOUSE cedex 9, France
[b]Bureau d'Enquêtes et d'Analyses (B.E.A), Aéroport du Bourget, 93352 Le Bourget cedex, France
[c]Laboratoire d'Acoustique de l'Université Toulouse-le Mirail, Université Toulouse-le Mirail, 5 allées Antonio Machado, 31058 Toulouse cedex 9, France
robert.ruiz@univ-tlse2.fr

Human voice carries non-linguistic information about emotion, fatigue, stress, truth, psychological illnesses etc. The proofs of this are well-established nowadays. In real-life situations, in laboratory conditions and from a cross-cultural point of view, the speaker's psycho-physiological disorders induce vocal modifications. Many acoustic parameters are measured. They belong to the dynamic and spectral planes. Phase space is also involved. Amplitude modulation is one of them. Unlike prosody and vocal quality features, this has not been widely studied. In this paper, a method for estimation vowel glottal pulses amplitude modulations is proposed. After pulse detection, a sinusoidal fit is applied leading to an estimate of the amplitude modulation frequency. The method is applied to experience sleep inertia effects on the voice. A pilot is suddenly awakened to undertake aeronautical psychomotor tasks. Results show the existence of an amplitude modulation. Their validity is based on determination coefficient measurements taking into account the number of pitch periods. Additionally, shimmer measurements show an increase after awakening. It can thus be concluded that sleep inertia has an effect on vowels uttered by the pilot.

1 Introduction

Many acoustic parameters are measured to study vocal effects of psychophysiological disorders on vowel signals. A lot of them are issued from the dynamic plane: pitch periods, jitter, shimmer, duration and prosody features; Other ones from the spectral plane (center of gravity and various energy ratios for example) [1 to 5] and from the phase space (maximal Lyapunov exponent) [6, 7, 13]. Original ones are also introduced [8, 10, 11, 12, 14, 15].

In this acoustic parameters large set, frequency and amplitude local time variations are measured by jitter and shimmer. But, these parameters do not provide information about the existence of signal modulations.

Frequency modulation, has been widely studied, especially for lying detection [9]. For amplitude modulation few studies exist [16, 17]. The first one is a device patent for measuring signal amplitude excursions. The second one shows significant variations of the amplitude modulation frequency for individuals with neurological diseases comparing to subjects producing vibrato (and not tremolo). Furthermore, they produce sustained vowel phonation and not natural speech.

In this paper, a simple detection method is proposed for segmented vowels. After pulses detection, a sinusoidal fit is applied to consecutive amplitude values of peak signals. Therefore, an estimation of amplitude modulation frequency is possible. The method is then applied to the study of sleep inertia effects on voice.

2 Amplitude Modulation of Vowel Glottal Pulses

Examination of vowel signals shows that peaks amplitude at the beginning of each vibration cycle can change during an utterance. These variations are not necessarily linked to a speaker's disorder. They occur in a neutral state.

The aim of the study is to determine if the shape drawn by their consecutive values can be a sine one, i.e if an amplitude modulation exists.

Figure 1 is an example of a digital amplitude-time display of a vowel uttered for the application experiment on sleep inertia. Circles are the amplitude peaks of the glottal pulses. A specific Matlab® program detects them to obtain successive pith period lengths.

Their inverse ratio gives the fundamental frequency F_0 of each cycle to compute jitter (1,29 Hz) and mean F_0 (153 Hz). From these successive amplitudes shimmer is also obtained (3,76 dB).

But one can notice that their chronological variation seems to have a sinusoidal aspect.

It is called amplitude modulation of glottal pulses. The purpose is to estimate its frequency modulation.

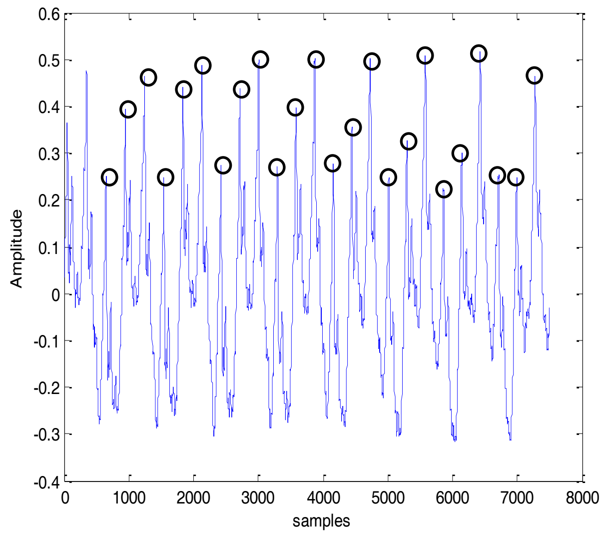

Figure 1 : Amplitude-time display of a [a] (sampling frequency 44,1 kHz). Circles are the amplitude peaks of the glottal pulses.

Before fitting, the set of consecutive glottal pulses amplitudes (circles of Figure 1) is replaced by the one of the differences between each amplitude and the mean value of the first set. Then, a sine fit is done on these normalized values by Matlab® (curve fitting toolbox).

Figure 2 shows both the sine regression and the broken line of consecutive normalized values. The amplitude modulation frequency is 50,1 Hz. There is no relationship with the frequency of the electricity power signal. In other experiments different values are obtained.

Because a fit can always be made, it is important to assess its validity to conclude the existence of an amplitude modulation. This is done by computing the determination coefficient R^2 between the set of fitted normalized amplitude values and the set of raw ones.

A R^2 close to one indicates that the regression line almost perfectly fits the data. If it is so, the amplitude modulation exists. Its frequency is the one of the sine function of the fit.

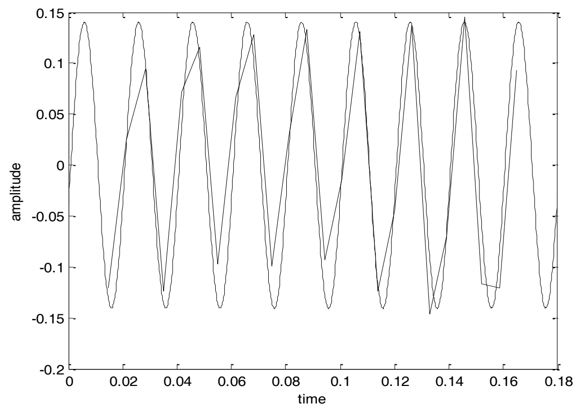

Figure 2: Broken line of the differences between each amplitude glottal pulse (circles of Figure 1) and the mean value – Sine fit.

It is the case for the vowel of Figure 1: $R^2 = 0,93$ (from Figure 2).

Goodness of the fit is the indicator of an amplitude modulation. Only values very close to 1 can allow to conclude (or values near 0).

Quality of the measurement is linked to the signal duration. Greater is the number of pith periods, better is the result. For short vowel signals, (4 cycles or less), an amplitude modulation frequency can be measured but the confidence degree is lower than with longer signals (for the same R^2). Indeed, it is easier to find a good sine fit on 4 values than on 20.

For that reason, the number of periodic patterns or cycles is always associated to the R^2 and frequency values of the fit.

When a set of vowels is analyzed, it is recommended to study the relationship between the amplitude modulation frequency (or R^2) and the number of cycles. This can be done by the drawing of the two quantities in a two-dimensional figure (possibly with the calculation of the corresponding correlation coefficient) (Figures 5 and 6).

3 Sleep Inertia Experiment

No assumption is suggested on the existence of an amplitude modulation in a neutral or in an upset state of the speaker or in both.

The topic of sleep inertia is important in aeronautics. During long flights, pilots sometimes decide to sleep alternately. They can also succumb to short periods of drowsiness. If a technical problem occurs, like an alarm in the cockpit, they have to react as rapidly as possible not only to evaluate the issue but also to solve it. The question that arises is whether ability to concentrate and work efficiency are always at an optimum level.

The purpose of the experiment was to create such a situation in the laboratory in order to record speech before and after a sudden awakening.

Some physiological variations are known (like heart rhythm). Voice modifications are also possible. Because of the non-invasive nature of acoustic measurements for pilots who are always recorded, such research can be of a great interest.

3.1 Experimental Conditions

The experiment is conducted in a French hospital (sleep disorders department). The speaker is a pilot. It is asked to read the same five sentences:

- upon its arrival at the lab (11 AM) (recording 1),
- after lunch (2 PM) (recording 2) and,
- after beiing suddenly awakened by a very powerful light (3 PM) (recording 3). Indeed, he has been beforehand taken to a quiet room where his sleep was monitored and was awakened a few minutes after falling asleep. Therefore, he performs various tasks on a computer, similar to flight ones, and reads the five sentences.

They are from aeronautical terminology and all contain an airplane registration ("Bravo, Victor, Charlie" for example). For example, one sentence is: "*Bravo, Victor, Charlie montez au niveau deux cinq zero*" which means "*Bravo Victor Charlie climb level two five zero*". The other four are the same type.

Sound recordings are made with a headset proximity microphone to maintain constant distance between the lips and the transducer for all head movements (AKG C555L). The recorder is a Sony Pro D.A.T with a sampling frequency of 44,1 kHz and a 16 bit resolution.

Vowels are segmented with an audio editing software so as to only keep quasi-stationary part of the signal. The attack, the decay and the transition with adjacent phonemes are suppressed. The monophtongs are analyzed during their stable period to measure the possible glottal pulses amplitude modulations.

The phoneme corpus counts 108 vowels: 37 in recording 1, 30 in recording 2 and 41 in recording 3.

3.2 Results

Amplitude modulation frequency f_{mod} is measured without any assumption and constraint about values of R^2 and number of cycles. Results are presented for all the vowels segmented in the five sentences uttered during the three recordings.

Figure 3 shows f_{mof} values. Most of them are close to 50 Hz. For the two neutral periods (recordings 1 and 2) about half of them are less than 50 Hz. But only 9 are in that case after the awakening (22% of the total number in recording 3 instead of 46% in the two first recordings).

The greater values of R^2 all belong to recording 3 (Figure 4). Most of them are very close to 1 which means that an amplitude modulation is present in the vowel signal. For recordings 1 and 2, amplitude modulation is not so obviously detected (excepting for three or four vowels). R^2 is not very close to one. So, amplitude modulation is not clearly present.

Sleep inertia seems to have an effect on vowels.

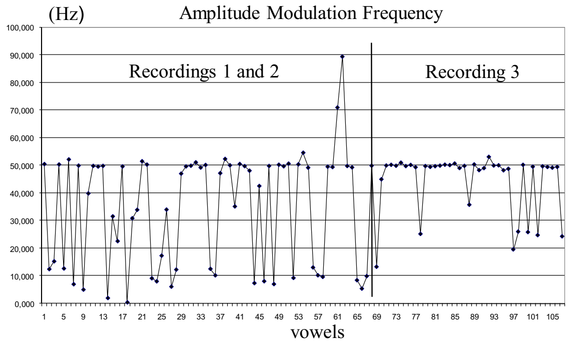

Figure 3: Amplitude modulation frequency f_{mod} for the vowels analyzed during the three recordings.

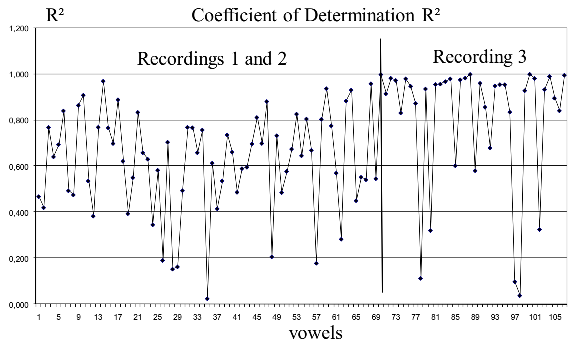

Figure 4: Coefficient of determination R^2 for the vowels analyzed during the three recordings.

To conclude on the existence or not of the modulation, the question that arises is whether there is a link between total cycle number N of the vowel and f_{mod} or R^2. For example, are lower values of R^2 due to a small number of pitch periods? Are higher f_{mod} values obtained for the greatest number of cycles?

In Figures 5 and 6, f_{mod} and R^2 are plotted for each vowel to answer these questions.

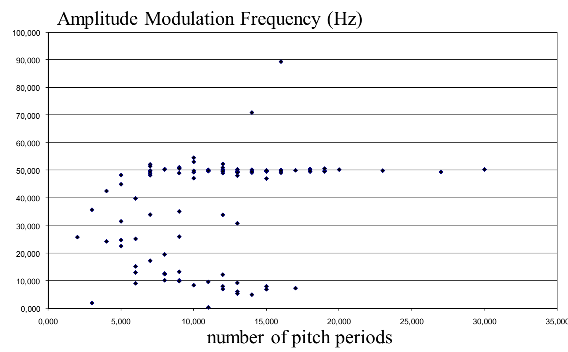

Figure 5: Amplitude modulation frequency and number of cycles for each vowel analyzed.

Examination of Figure 5 shows that the 50 Hz values of f_{mod} are obtained for numbers of pitch periods N varying from 7 to 30. But when the number of cycles is too small (less than 7), f_{mod} values of 50 Hz are not reached. It is the opposite when N is important (greater than 17).

It is recommended to have enough periods to guarantee measurement validity. Below four or five periods only a beginning of modulation can be found. Indeed, the vowel signal amplitude is modulated at low frequency.

Figure 6 shows that the R^2 values which are very close to one can be found from 2 to 14 pitch periods.

These points are those of vowels belonging to the third recording.

The greatest numbers of cycles do not necessarily lead to a good sine fit. The scatter of points indicates that there is no relationship between R^2 and the number N of pitch periods.

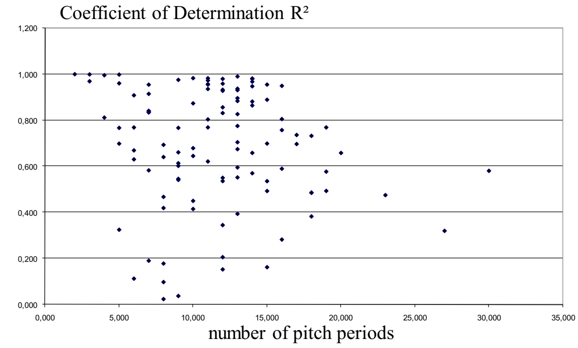

Figure 6: Coefficient of determination R^2 and number of cycles for each vowel.

Finally, it seems that an effect of sleep inertia on vowels is the appearance of an amplitude modulation. Its frequency is here 50 Hz and does not depend on the number of pitch periods.

Additional results follow.

The mean fundamental frequency is 132 Hz (standard deviation $\sigma = 32$ Hz): 130 Hz for recording 1 and 2 ($\sigma = 25$ Hz), 137 Hz for recording 3 ($\sigma = 41$ Hz). A greater dispersion of pitch values could also be a result of sleep inertia.

Shimmer S (in dB) which explores amplitude variations of the glottal pulses peak amplitude is measured.

$$S(dB) = \frac{1}{N-1}\sum_{i=1}^{N-1}\left|20\log_{10}\left(\frac{A(i)}{A(i+1)}\right)\right| \qquad (1)$$

A(i) is the amplitude of the peak *i*, *N* is the total number of vowel periods.

Figure 7 shows the variations of S for all the vowels analyzed in the three recordings.

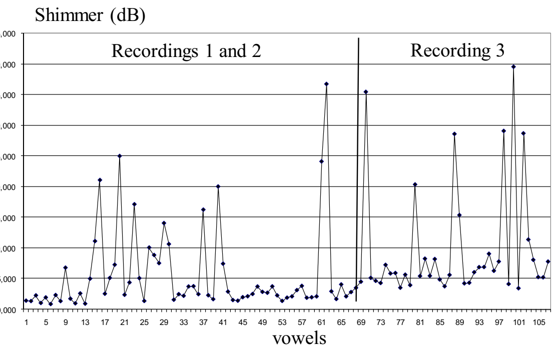

Figure 7: Shimmer for all vowels analyzed.

A statistically nonsignificant trend of S exists in recording 3: S = 5,5 dB for recordings 1 and 2 (σ = 6,9 dB) and S = 9,75 dB for recording 3 (σ = 9,4 dB).

For this speaker, sleep inertia seems to introduce a shimmer increase, an amplitude modulation of the glottal pulses, an increase of the fundamental frequency F_0 dispersion but no F_0 variation.

4 Conclusion

Amplitude modulation of glottal peaks is not audible. It is different from an audible tremolo for sustained vowels. With tremolo, the periodic increasing and decreasing of amplitude is observed on signal frames of several pitch periods. Here, the modulation only concerns glottal peaks (Figure 1) on short durations.

The problem is then to have a sufficient number of pitch periods in the segmented vowel. Measurements of this amplitude modulation with high speaking rate can become difficult. Even if the application to sleep inertia has shown that only a few periods (4 or 5) are necessary to compute f_{mod}, it is suitable to work with longer vowels. Indeed it is relatively easy to fit 4 or 5 values with a sine if they have quasi regular different amplitudes. More complicated is to obtain a good fit with more values. The quality of the measurement is closely related to the speaking rate even with high values of R^2.

The value of the amplitude modulation frequency is 50 Hz in this application to sleep inertia. This value, without any relationship with the frequency of the electricity power signal, is indicative and cannot be generalized: only one speaker has done the experiment.

The purpose was to test the measurement method. It is not sure that amplitude modulation exists for other speakers after awakening, and that the 50 Hz value remains the same. Complementary investigations are needed to study this amplitude modulation for neutral utterances and to confirm its presence in an unusual psychophysiological state of the speaker.

Acknowledgments

This work is supported by the B.E.A ("Bureau d'Enquêtes et d'Analyse pour la sécurité de l'aviation civile").

The authors wish to thank the " Laboratoire d'Anthropologie Appliquée" of the University Paris 5 and the "Laboratoire d'Exploration du Sommeil" of the hospital Henri Mondor which agreed voice recordings in their medical experimentations on sleep inertia. The pilot is also thanked for allowing his voice to be used.

References

[1] Lieberman, P., "Perturbations in vocal pitch", J. Acoust. Soc. Am., 33(5):597:603, 1961.

[2] Brenner, M., and Shipp, T., "Voice stress analysis", NASA N88-23395 report, 1987.

[3] Proceedings of Tutorial and Research Workshop on Speech under Stress, ESCA-NATO, Lisbon, Portugal, 14-15 Sept. 1995.

[4] Hansen, J.H.L., Swail, C., South, A.J, Moore, R.K., Steeneken, H., Cupples, E.J., Anderson, T., Vloeberghs, C.R.A, Trancoso, I., and Verlinde, P., "The impact of speech under stress on military speech technology", NATO project 4 report, ISBN 92-837-1027-4, March 2000.

[5] Devillers, L., Vidrascu, L., "Real life emotions detection with lexical and paralinguistic cues on human-human call center dialogs", Interspeech ICSLP, 801-804, Pittsburgh, 2006.

[6] Herzel, H., Berry, D., Titze, I.R., and Saleh: M., "Analysis of vocal disorders with methods from nonlinear dynamics", J.Speech and Hearing Disorders 37:1008-1019, 1994.

[7] Jiang, J.J., Zhang, Y., and Mc Gilligan, C., "Chaos in voice, from modelling to measurement", Journal of Voice, 20(1):2-17, 2006.

[8] Simonov, P.V., and Frolov, M.V., "Utilization of human voice for estimation of man's emotional stress and state of attention", Aerospace Med., 44(3):256-258, 1973.

[9] Older, H.J., and Jenney, L.L., "Psychological stress measurement through voice output analysis", NASA CR 141723, N75-19960 report, March 1975.

[10] Kuroda, I., Fujiwara, O., Okamura, N., and Utsuki, N., "Method for determining pilot stress through analysis of voice communication", Aviat. Space Environ. Med., 47(5):528-533, 1976.

[11] Shiomi, K., "Fatigue and Drowsiness Predictor for Pilots and Air Traffic Controllers", Proc. of 45th Annual ATCA Conference, Oct. 2000.

[12] Shiomi, K., Sato, M., and Sawa, A., and Suzuki, "Experimental Results of Measuring Human Fatigue by Utilizing Uttered Voice Processing", Proc. of IEEE-SMC 2008, P557, Singapore, 2008.

[13] Ruiz, R., Plantin de Hugues, P., and Legros, C., "Advanced Voice Analysis of Pilots to Detect Fatigue and Sleep Inertia", Acustica united with Acta Acustica, accepted for publication, 2010.

[14] Ruiz, R. and Legros, C. "The Cumulative Spectral Probability Diagram: Theory and Experiments", Acustica united with Acta Acustica, 2 (3):215-222, 1994.

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

[16] Fuller, F.H., Method and apparatus for phonation analysis leading to valid truth/lie decisions by spectral energy region comparison, United States patent N° 3,855,417, Dec. 17, 1974.

[17] Winholtz, W.S., Ramig, L.O., Vocal tremor analysis with the vocal demodulator, J. of Speech and Hearing Research, 35, pp. 562-573, 1992.