www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 2 Issue 12, Dec, 2013 Page No. 3439-3443

Effect of Aspirated Consonants on EMG Signals Generated in Zygomaticus Muscles

Himanshu Raina[1], Randhir Singh[2], Parveen Lehana[3]

M.Tech, Department ECE, Shri Sai College of Engineering and Technology, Pathankot, India[1]
HOD, Dept ECE, Shri Sai College of Engineering and Technology, Pathankot, India[2]
Associate Professor, Dept of Physics and Electronics, University of Jammu, India[3*]
pklehanajournals@gmail.com

Abstract: ***Speech communication is one of the simplest and reliable forms of communication between humans and it is affected by the behaviour and emotions of speakers. Electrical current generated on the facial muscles during speech production are represented by one of the forms of biomedical signal known as EMG. These signals results due to contraction and relaxation of the muscles, which are controlled by the nervous system. These signals from specific facial muscles are recorded for speech recognition and system automation. EMG signals are generally recorded using small surface electrodes placed near to each other. EMG activity is frequently recorded from specific muscles and plays a prominent role in the expression of elementary emotions and speech generation. The present research paper investigates the EMG patterns generated during the utterance of the unvoiced consonants. Six subjects in the age of 20-25 years were taken (three males and three females). Thirty eight vowel-consonant-vowel (VCV) syllables in Hindi were recorded along with the corresponding facial EMG signal. For each speaker, the means of log-spectral-distances (LSD) between the EMG signal of the VCVs and the reference EMG signal were computed. Analysis of the spectrograms and LSD showed that the EMD signals generated in the muscle vary with the subject and the VCV. Hence, for automatic decoding of the EMG signals, the system should be trained using both the variants.***

Keywords: EMG signals, speech, aspirated consonants, spectrogram.

1. Introduction

Acquisitions of electrical signal from any organ that represent a physical variable are known as biomedical signals. Acquired signals are usually function of time but desirable in terms of amplitude, frequency and phase. Electrical current generated in muscles due to contraction are the EMG signals [1]. These signals are complicated and controlled by nervous system. Anatomical and physical properties of the muscles significantly affect the EMG signals. EMG signals have application in the area of clinical diagnosis and biomedical based system design [2]. Facial electromyography (EMG) provides useful information regarding emotion, fatigue or feeling. Also Facial EMG signals are implemented in speech recognition system. EMG signals are generated in the facial muscles due to movement of lips, lifting the corners of the month, or opening the jaw. Anatomical view of the muscles shown in Fig. 1 responsible for the generation of facial EMG during speech production [3]. A general procedure to record the EMG signal can be divided into three stages; electrode selection (surface or needle), placement of electrodes, and signal conditioning setup. Inorder to minimize or eliminate possible noise all these stages must be handled carefully. Among these stages electrode is an important interface, which can pick up the biopotential or the current in the human body, between the human body and measurement instrument [4]. Needle electrodes are preferred in clinical test, as it provides selectively of muscles under observation and reduces noise also [5]. Half cell potential of electrode not only depends upon the type of the electrode but also material of the electrode influences it. In order to minimize the half cell potentials difference, electrodes of same material are used in EMG measurement.

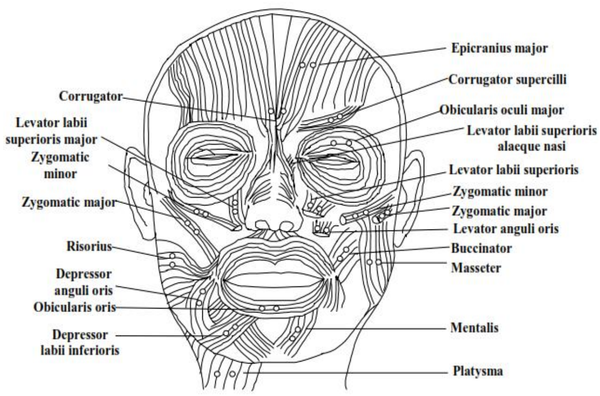

Figure 1: The location of facial muscles [6].

Facial EMG signal recorded during the human speech production mechanism may be implemented for speech recognition and also for designing an automated speech recognition based systems.

2. Human Speech Production

Set Speech signals are in the form periodic wave. Various forms of excitation methods are applied to produce voice, unvoiced etc speech signals. In voiced excitation, glottis open and closed periodically by air pressure and produced a periodic pulse which is in triangle shape. Fundamental frequency lies in the range from 80Hz to 350Hz for this type of excitation. Glottis is open and the air passes a narrow channel in the mouth or throat in unvoiced excitation. Unvoiced excitation results disturbances in the sound signal which generates noise [6-7]. Speech model generated by human is given in Fig. 2.

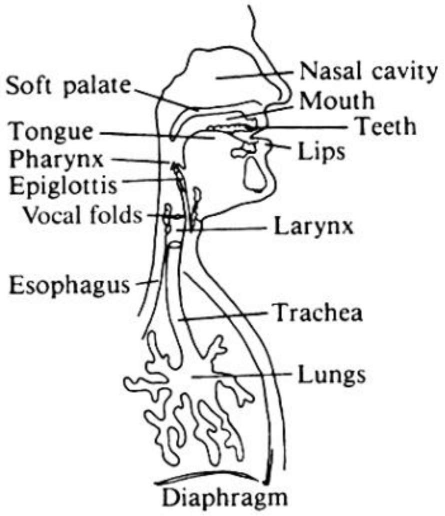

Figure 2: The human speech production organs [8].

Various artificial models are designed by many researchers [9-19]. In this section two artificial human vocal tract models are discussed. Umeda and Teranishi developed a simple human vocal tract device that simulated human speech acoustically shown in Fig. 3. It consists of moving thick plate strips, cross sectional areas are changed when these strips are inserted closely from one side. According to the configuration of the model various vowels and other sustained sounds are produced. Glottal signals are sent into glottis of the model and emitted from the mouth.

Figure 3: Artificial model of the human vocal tract by Umeda and Teranishi [18].

In Large lung model shown in Fig. 4, works when the knob pulled which is attached to the rubber membrane covering the bottom of the cavity to inflate two balloons acting as lungs. Trachea is simulated by Y-shaped tube which is attached to these balloons. A negative pressure in the air inside the thoracic cavity is created by lowering the diaphragm increasing the volume of the thoracic cavity. Air flows into the lungs to equalize the pressure inside the lungs with atmospheric pressure, simulating inhalation [20-23].

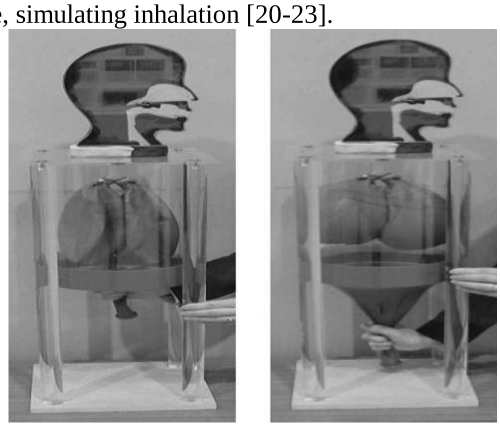

Figure 4: Large lung models with head shaped models are set on the top of these large lung models [22].

3. Methodology

Research work is carried out to investigate the EMG patterns generated during the utterance of the aspirated consonants. For investigating facial EMG patterns generated by uttering unvoiced consonants in VCV syllables, experiments were conducted with the six subjects (three males and three females) having age between 20-25 years. The speech signals and the corresponding EMG signals were recorded using a data acquisition system at the sampling frequency of 16 kHz and 16-bit quantization. Electrodes were placed at zygomaticus minor, zygomaticus major, and mentalispoints on the face. Block diagram of the experimental setup for recording is shown in Fig. 5. The recorded signals were segmented and labeled manually into separate files for each of the VCVs. The signals were analyzed using time-domain patterns, spectrograms, and mean log-spectral-distances. For computing LSD, the first VCV of each subject was taken as the reference.

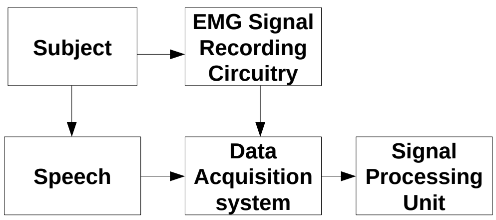

Figure 5: Block diagram for EMG signal acquisition.

4. Results and Conclusion

Investigations were carried out to evaluate the effect of utterance of aspirated consonants आखा and आछा, and generation of facial EMG during the production of these consonants.

Time-domain signals and the corresponding spectrograms of the speech and facial EMG signals for two unvoiced consonants आखा and आछा are shown in Fig. 6 to Fig. 11 for six subjects (Sp1 to Sp3 are male subjects and Sp4 to Sp6 are female subjects). Here the x-axis represents the normalized time and y-axis represents the normalized frequency. The signals and the corresponding spectrograms show that the signals generated vary across the subjects and the syllables.

Table I shows the means and standard deviations of unvoiced consonants आखा and आछा of all the six speakers.

Table 1: Mean and standard deviation of two voiced consonant.

Speaker	Aspirated consonant (आखा)		Aspirated consonant (आछा)	
	Mean	S.D.	Mean	S.D.
Sp1	7.16	3.51	9.39	2.51
Sp2	6.77	1	7.95	1.74
Sp3	5.33	1.3	5.26	1.18
Sp4	6.54	0.69	7.04	0.76
Sp5	5.93	1.2	6.84	1.01
Sp6	6.76	3.88	13.34	5.24

Analysis of the mean LSDs also suggests that the distances are a function of subject and VCV. Hence, training of the automated EMG recognition system needs both the EMG signals and the information regarding subjects.

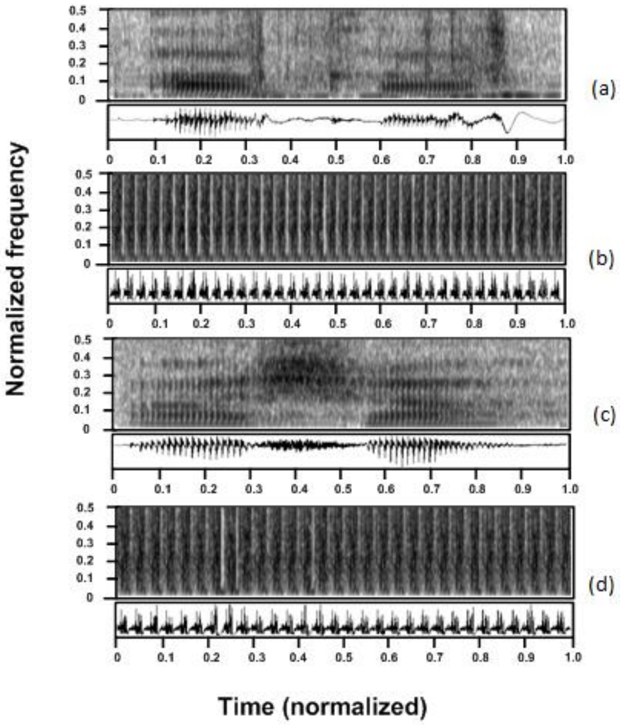

Figure 6: Spectrogram of aspirated consonant (a) Original speech signal (आखा) (b) Facial EMG signal (आखा) (c) Original speech signal (आछा) (d) Facial EMG signal (आछा) for Sp1.

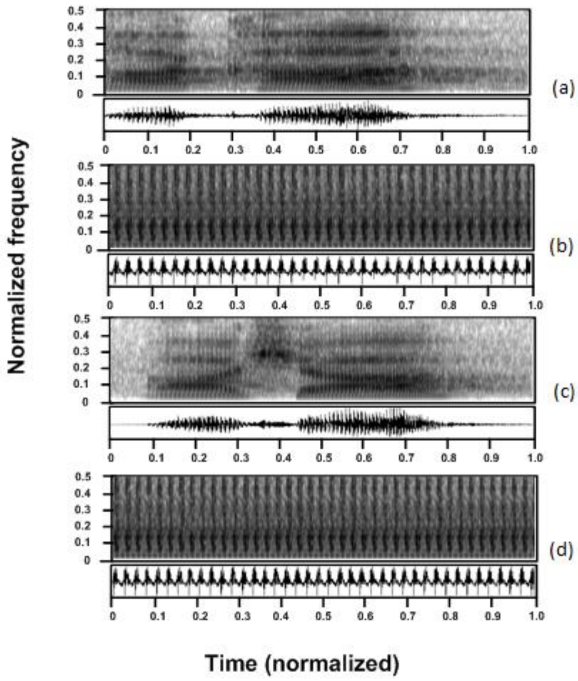

Figure 7: Spectrogram of aspirated consonant (a) Original speech signal (आखा) (b) Facial EMG signal (आखा) (c) Original speech signal (आछा) (d) Facial EMG signal (आछा) for Sp2.

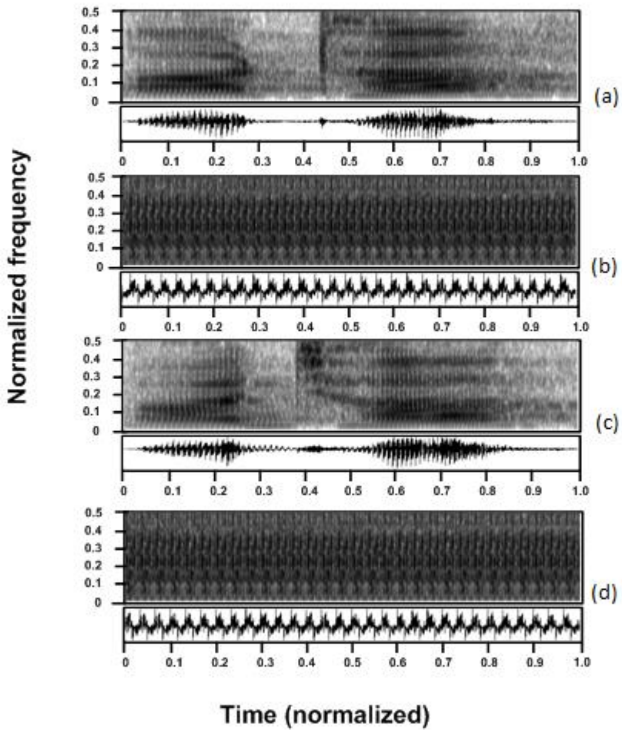

Figure 8: Spectrogram of aspirated consonant (a) Original speech signal (आखा) (b) Facial EMG signal (आखा) (c) Original speech signal (आछा) (d) Facial EMG signal (आछा) for Sp3.

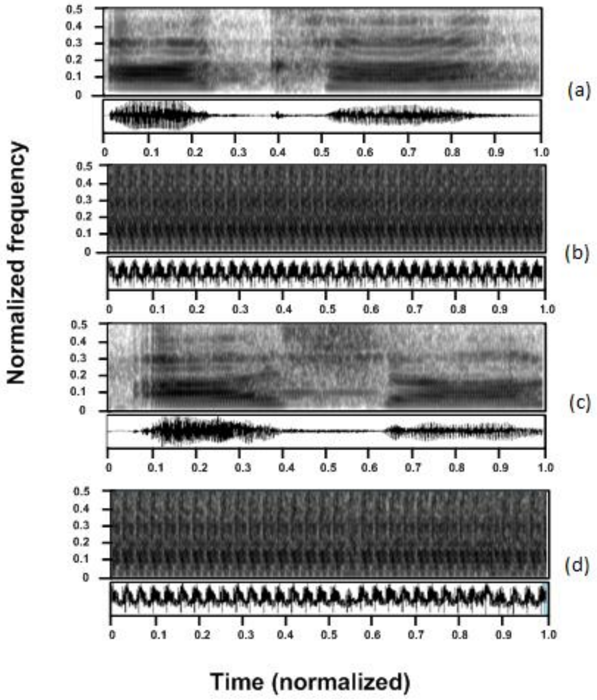

Figure 9: Spectrogram of aspirated consonant (a) Original speech signal (आखा) (b) Facial EMG signal (आखा) (c) Original speech signal (आछा) (d) Facial EMG signal (आछा) for Sp4.

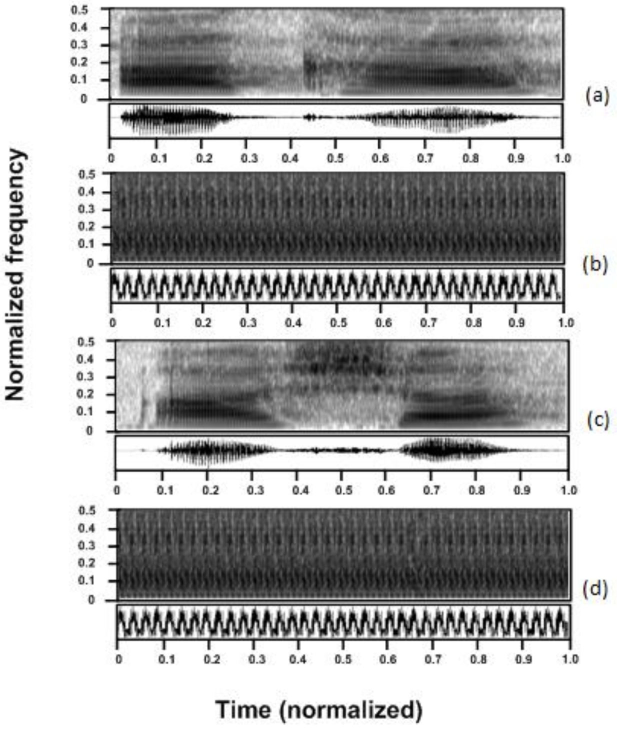

Figure 10: Spectrogram of aspirated consonant (a) Original speech signal (आखा) (b) Facial EMG signal (आखा) (c) Original speech signal (आछा) (d) Facial EMG signal (आछा) for Sp5.

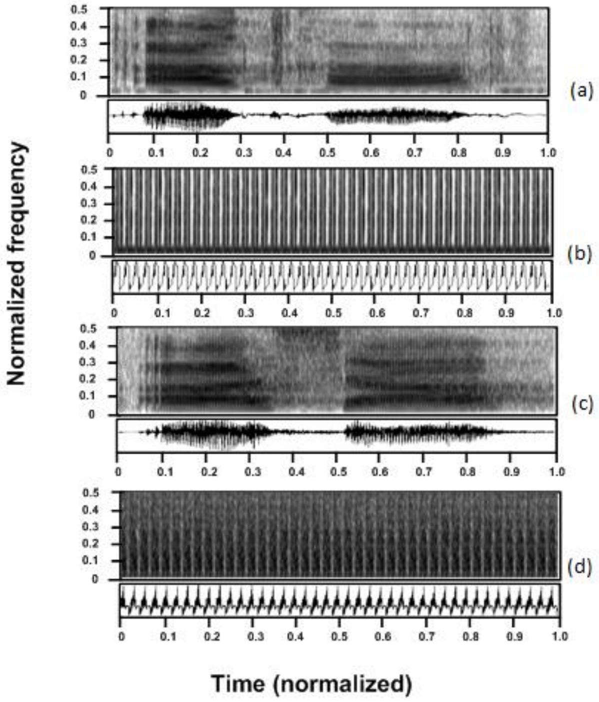

Figure 11: Spectrogram of aspirated consonant (a) Original speech signal (आखा) (b) Facial EMG signal (आखा) (c) Original speech signal (आछा) (d) Facial EMG signal (आछा) for Sp6.

References

[1] M. B. I. Reaz, M. S. Hussain and F. M. Yasin, "Techniques of EMG signal analysis detection, processing, classification and applications," Biological Procedures Online, vol.8, 2006 pp. 11-35.

[2] J. V Basmajian and C. J d. Luca. Muscles Alive The Functions Revealed by Electromyography, The Williams & Wilkins Company Baltimore, 1985.

[3] B. G .Lapatki, D. F. Stegeman, and I. E. Jonas, "A surface EMG electrode for the simultaneous observation of multiple facial muscles," Journal of Neuroscience Methods, vol1.23,2003, pp.117-128.

[4] M. Bakke, "Mandibular elevator muscle physiology action, and effect of dental occlusion," European Journal of Oral Sciences, vol. 101, pp. 314-331, 1993.

[5] C. J. D Luca, "The use of surface electromyography in biomechanics," Journal of Applied Biomechanics, 1997.

[6] R. Mugitani and S. Hiroya, "Developed of vocal tract and acoustic features in children," Acoustic Science & Technology, 2012.

[7] T. D. Rossing, F. R. Moore and P. A. Wheeler. The Science of Sound. Addison Wesley, 3rd edition, 2002.

[8] W. de, M. V. Boxtel, A. Zaalberg, R. Goudena, P.P, and W. Matthys, "Facial EMG responses to dynamic emotional facial expressions in boys with disruptive behavior disorders," Journal of Psychiatric Research, vol.40, 2006, pp.112-121.

[9] P. Ekman, and W.V. Friesen, "Facial Action Coding System (FACS). A technique for the measurement of facial action," Consulting Psychologists Press Palo Alto, CA, 1978.

[10] P. Elman, J. C. Hager, and W.V Friesen, “The symmetry of emotional and deliberate facial actions” Psychophysiology, 18 1981, pp. 101-106.

[11] B. v Anton, “Facial EMG as a Tool for Inferring Affective States,” Proceeding of Measuring Behavior, Netherlands August, pp. 24-27, 2010.

[12] J. H. Hermie, F. Bart, D. K .Catherine and R. Gunter, “Development of recommendations for SEMG sensors and sensor placement procedures,” Journal of Electromyography and Kinesiology, vol.10, 2000 pp. 361-374.

[13] S. Deketelaere, O. Deroo, and T. Dutoit, “Speech processing for communications what’s new?,” Copernic Ave, Initialis scientific park.

[14] B. H. Juang, and L. R. Rabiner, “ Hidden markov models for speech recognition,” Technometrics, AT&T bell laboratories, Murray hill, vol. 33, pp. 251-272, 1991.

[15] S. E. Masri, X. Pelorson, P. Saguet, and P. Badin, “Vocal tract acoustics using the transmission line matrix (TLM) method,” in Proc. of the Fourth International Conference on Spoken Language Processing, Philadelphia, pp. 953–956, 1996.

[16] C. Lu, T. Nakai, and H. Suzuki, “Finite element simulation of sound transmission in vocal tract,” Journal of Acoustic Society Japan, 1993 pp. 2577–258.

[17] N. Umeda, and R. Teranishi, “Phonemic feature and vocal feature: Synthesis of speech sounds, using an acoustic model of vocal tract*,”* Journal of Acoustic Society Japan, 1966 pp. 195-203.

[18] T. Arai, “The replication of Chiba and Kajiyama's mechanical models of the human vocal cavity,” Journal of the Phonetic Society of Japan, 2001 pp.31-38.

[19] T. Arai, E. Maeda, N. Saika, and Y. Murahara, “Physical models of the human vocal tract as tools for education in acoustics,” in Proc. of the First Pan-American/Iberian Meeting on Acoustics in Cancun, 2002.

[20] T. Lander and T. Arai, “Using Arais vocal tract models for education in Phonetics,” in Proc. of International Conference on Spoken Language Processing, Barcelona, , pp .317-320, 2003.